Charles University
Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Optimal portfolio selection under Expected Shortfall optimisation with Random Matrix Theory denoising

Author: **Bc. Jan Šíla, MSc.**
Supervisor: **PhDr. Boril Šopov, MSc., LL.M.**
Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 5, 2018

Signature

Acknowledgments

Věnování této práce si více nezaslouží nikdo jiný, než moji rodiče a to za celoživotní podporu ve studiích akademických, i životních. Za to, že se od školky starali o to, abych se mohl soutředit na věci, jež mne zajímaly, plánovat cestu do vesmíru, a nemusel chodit tak často s odpadky. Zejména pak děkuji za nákupy a skaldování mnoha encyklopedií a knih, jejichž obsah si nyní jen zřídka pamatuji. Za to vám patří velký obdiv a tato sekce práce v češtině, neboť svou angličtinu pečlivě zanedbáváte.

Chtěl bych také poděkovat svým dvěma sestrám, za soustavné udržování krevního tlaku na vysoké úrovni; tolik vhodné ke studiu a dospívání.

Bibliographic entry

ŠÍLA, J. (2018): "Optimal portfolio selection under Expected Shortfall optimisation with Random Matrix Theory denoising" (Unpublished diploma thesis). Charles University in Prague

Extent: 91 832 characters including spaces

Abstract

This thesis challenges several concepts in finance. Firstly, it is the Markowitz's solution to the portfolio problem. It introduces a new method which denoises the covariance matrix – the cornerstone of the portfolio management. Random Matrix Theory originates in particle physics and was recently introduced to finance as the intersection between economics and natural sciences has widened over the past couple of years.

Often discussed Efficient Market Hypothesis is opposed by adopting the assumption, that financial returns are driven by Paretian distributions, instead of Gaussian ones, as conjured by Mandelbrot some 50 years ago.

The portfolio selection is set in a framework, where Expected Shortfall replaces the standard deviation as the risk measure. Therefore, direct optimisation of the portfolio is implemented to be compared with the performance of the classical solution and its denoised counterpart. The results are evaluated in a controlled environment of Monte Carlo simulation as well as using empirical data from S&P 500 constituents.

JEL Classification C46, C53, C58, D81, G11, G32

Keywords Portfolio management, Random Matrix Theory, Lévy–stable distributions, Expected Shortfall

Author's e-mail `61558906@fsv.cuni.cz`

Supervisor's e-mail `boril.sopov@gmail.com`

Abstrakt

Tato práce se zabývá několika koncepty moderních matematických financí. Předně je to teorie portfolia, jak ji v 50. letech předchozího století formuloval Harry Markowitz. Práce se týká nové metody, kterou do moderního finančního světa přinesla částicová fyzika před několika lety. Random Matrix Theory (Teorie náhodných matic) si klade za cíl vyčistit kovarianční matici od šumu, který při klasickém Pearsonovském odhadu nutně vzniká. Tím vzniká hypotéza, zdali takto upravené kovarianční matice povedou k lepším investičním rozhodnutím, zejména vzhledem k risku daného portfolia.

Inspirována Benoit Mandelbrotem si tato práce osvojuje Stabilní rozdělení, jakožto předpokládané rozdělení pro finanční instrumenty, což je v přímé opozici s Efficient Market Hypothesis (Hypotéza efektivních trhů).

Teorie portfolia v tomto případě odhlíží od míry risku měřené pomocí standardní odchylky, přičemž se zabývá mírou Expected Shortfall. Práce porovnává obě metody s přímou optimalizací portfolia dle minimalizace právě Expected Shortfallu.

Výsledky jsou vyhodnoceny za použití simulační metody Monte Carlo, stejně jako empirických dat z akciových titulů zastoupených v indexu S&P 500.

Klasifikace JEL	C46, C53, C58, D81, G11, G32
Klíčová slova	Teorie portfolia, Random Matrix Theory, Stabilní rozdělení, Expected Shortfall

E-mail autora	61558906@fsv.cuni.cz
E-mail vedoucího práce	boril.sopov@gmail.com

Contents

List of Tables

List of Figures

Acronyms

RMT Random Matrix Theory, or denoised covariance matrix

GMB Geometric Brownian Motion

MLE Maximum Likelihood Estimate

VaR Value at Risk

ES Expected shortfall of a portfolio

cVaR conditional Value at Risk – synonym to ES

C Pearsonian covariance matrix estimate

Q Noise proxy for portfolio - number of observations to number of assets

PSD Positive Semi-Definit matrix

Master's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University

Author:	**Bc. Jan Šíla, MSc.**	Supervisor:	PhDr. Boril Šopov, MSc., LL.M.
E-mail:	61558906@fsv.cuni.cz	E-mail:	boril.sopov@gmail.com
Phone:	603941008	Phone:	606637608
Specialization:	FFM&B	Defense Planned:	February 2018

Notes: The proposal should be 2-3 pages long. Save it as "yoursurname_proposal.doc" and send it to mejstrik@fsv.cuni.cz, tomas.havranek@fsv.cuni.cz, and zuzana.havrankova@fsv.cuni.cz. Subject of the e-mail must be: "JEM001 Proposal (Yoursurname)".

Proposed Topic:

Minimization of Expected shortfall under Random Matrix Theory and stable distributions

Motivation:

Much of the intellectual energy in finance of the 20th century was spent on the problem of optimal portfolio selection. Historically, the cornerstone of the solution, presented by Markowitz in 1952, has been the variance-covariance matrix – a measure of co-movement of financial time series. However, there are inherent weaknesses in this approach that have been challenged since the introduction of methods originating in other fields – Physics and signal processing - into Economics.

Random Matrix Theory is one of the most often scrutinised theories about its ability to offer better estimates of the 'true' covariances. This work shall implement RMT denoising method, in order to improve the solution of the portfolio problem, particularly with respect to the risk.

Furthermore, the original problem setup minimises the risk expressed as the standard deviation of the returns of the portfolio, given a level of return. This thesis looks into the problem of minimising the Expected shortfall of the portfolio instead.

With the advent of Basel III accord, the mainstream measure of risk shifts from Value at Risk to a more sophisticated Expected shortfall. Also, with soaring power of today's standard hardware, we are able to employ more computationally exhaustive methods in portfolio optimization and risk management as such.

Instead of assuming that financial data follow Gaussian laws, semi-closed form of alpha-Stable Lévy distributions can be taken into account. Unlike in the world of Normal distributions, Stable distributions have infinite variance and are a generalization of several important distributions. Hence modelling the conditional Value at Risk, we can account for the fat tails more precisely than in the usual framework.

This work will test the optimization method on both simulated data, with known covariance structure, as well as real data to empirically test the hypotheses described below.

Hypotheses:

1. Hypothesis #1: RMT will be tested to be a better approximation of the data-generating covariance structure that Pearsonian covariance matrix.
2. Hypothesis #2: Random Matrix Theory denoising of the covariance matrix will result in portfolios with lower Expected shortfall when portfolio is optimised with respect to it.
3. Hypothesis #3: The ratio of number of observations and number of stocks in the portfolio will be significant to the performance of the method.

Methodology:

The method of Random Matrix Theory has not been inspected in quantitative finance journals over the past decade very frequently. It has found its core authors and there have been some interesting empirical results.

The methodology is not however very clear on the method of clipping the noisy eigenvalues of the covariance matrix. Hence it will be inspected and implemented from scratch to avoid any possible flaws in reusing external code.

The semi-closed solution to the ES minimisation will be presented with adequate mathematical background and implemented in respective software as well.

The stable distributions will be studied and for there is quite a lot of confusion in the literature, and particularly implementation, the work will follow Nolan (2003, 2012). A Maximum Likelihood Estimate of the stable distribution fits to the time series will be inspected, despite its computational exhaustiveness.

Lastly, the thesis shall follow and verify the method of calculating the expected shortfall for Stable distributions due to Stoyanov (2006).

The author is expecting to deliver its own implementation of the methods in R programming language or Python.

Expected Contribution:

The work will summarise the research in Random Matrix Theory, Stable distributions and modern Portfolio Theory.

It attempts to firstly verify the relevance of RMT with simulating multivariate financial time series and comparing it to Pearsonian estimates using Kullback-Leibler measure. Thus validating and replicating Macenko (1967) results. Secondly, it will inspect, the effect of using this method on creating portfolios with respect to the Expected Shortfall of the resulting portfolio – it will partially replicate results of Laloux (2000), Edelman (2005) and Bouchaud (2009). This will be compared to portfolios calculated as Global Minimum Variance from the usual Markowitz's framework.

Also, the thesis will implement and replicate core results for ES estimation by Stoyanov (2006).

Moreover, the thesis will offer discussion on using Stable distributions in finance and clarify the confusion behind various parametrizations – Nolan(2012). It will implement a library that calculates the stable density as in Nolan (2012).

Moreover, it will implement an MLE estimate engine for fitting stable distributions. Furthermore, it shall evaluate Expected Shortfalls for various portfolios and their parameters such as rebalancing, under parametric assumptions as well as actual historical performances.

Outline:

Chapter 1 - Introduction
Chapter 2 – Literature Review – almost finished
Chapter 3 – Methodology – mostly done
Chapter 4 – Data and results
Chapter 5 - Conclusion

Core Bibliography:

Bouchaud, J.-P. & M. Potters (2009): "Financial applications of random matrix theory: a short review." arXiv preprint arXiv:0910.1205 .

Edelman, A. & N. R. Rao (2005): "Random matrix theory." Acta Numerica 14: pp. 233–297.

Gnedenko, B. V. & A. N. Kolmogorov (1954): Limit distributions for sums of independent random variables. Addison-Wesley Mathematics Series. Cambridge, MA: Addison-Wesley. Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob. MR:0062975. Zbl:0056.36001.

Laloux, L., P. Cizeau, M. Potters, & J.-P. Bouchaud (2000): "Random matrix theory and financial correlations." International Journal of Theoret- ical and Applied Finance 3(03): pp. 391–397.

Marčenko, V. A. & L. A. Pastur (1967): "Distribution of eigenvalues for some sets of random matrices." Mathematics of the USSR-Sbornik 1(4): p. 457.

Nolan, J. P. (2003): "Modeling financial data with stable distributions." Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book 1: pp. 105–130.

Nolan, J. P. (2012): Stable distributions, volume 1177108605. ISBN.

Stoyanov, S. V., G. Samorodnitsky, S. Rachev, & S. Ortobelli Lozza (2006): "Computing the portfolio conditional value-at-risk in the alpha- stable case." Probability and Mathematical Statistics 26(1): pp. 1–22.

Zolotarev, V. M. (1986): One-dimensional stable distributions, volume 65. American Mathematical Soc.

Author **Supervisor**

Chapter 1

Introduction

Since the beginning of the last century, finance practitioners were convinced that the markets are immeasurable. Or rather, effort to analyse and predict them in order to obtain profits, is doomed to fail. This idea originated with Bachelier's thesis on speculation on options markets and laid the foundation stone of the Efficient Market Hypothesis.

Finance and its methods have substantially evolved, particularly since the turn of the century, towards a more quantitative side of things. This greatly enhanced its ability to forecast and analyse the markets, also in real–time. Finance has received contributions from exact natural sciences, on top of the original economics. Nowadays, finance consist of a mix of methods from stochastic calculus, data mining or signal processing, all of which are widely used in practice as well as academia. Along with a massive growth in computational power of conventional computers, modern time researchers can take advantage of methods solving tasks in complexity quite unimaginable just a couple of decades back. Yet some core economic ideas remain quite unchallenged, which raises a question whether or not they are still relevant.

Plenty of the intellectual resources invested in finance of the previous century were spent on portfolio theory and portfolio risk management. However, since the pioneering work of Markowitz in 1952, the setup has not seen much innovation.

This work is concerned with the cornerstone of the conventional portfolio theory – covariance of assets' returns. It inspects effects, and possibly the benefits, of denoising the estimates of the covariance matrix to the portfolio performance. Thus, it is testing whether or not the resulting portfolios can be less risky, as they would have more accurate inputs. The considered method is Random Matrix Theory, which originated in particle physics and its study of nuclear resonance with random matrices. Validity of this approach is tested under controlled conditions as well as using empirical data.

As a risk measure, this work replaces the original objective of minimisation of standard deviation with the Expected Shortfall. This is motivated by the recent regulatory departure from the Value At Risk to its more sophisticated version known as conditional VaR, or Expected Shortfall. It has more suitable mathematical properties, as unlike the Value at Risk, it satisfies the coherency axioms.

Moreover, this work disregards the assumptions that financial series originate in Gaussian world. Rather, it implements more general Lévy α-stable distributions. This family of distributions is proposed by Mandelbrot to oppose the Efficient Market Hypothesis. Also, a semi–closed solution to calculating the Expected shortfall under Stable distributions, was proposed only recently by Stoyanov. Hence, results presented herein are one of the few applications of this important formula.

Moreover, this work directly compares predictions by RMT and Markowitz's framework, with portfolios optimised with respect to their Expected Shortfall. Furthermore, the theoretical part presents steps how to execute such calculations more effectively than is standard in the scripting language R used in this thesis.

The results of the empirical part are re–evaluated using Monte Carlo simulations to obtain more representative conclusions. Neither set of results suggests, that Markowitz's solution performs consistently worse that the more modern counterparts. Random Matrix Theory validates its effectiveness in situations with high level of noise and minimisation of the Expected shortfall proves less flexible in adverse market conditions than the other two methods.

The thesis is structured as follows: Chapter 2 presents literature review of the main topics of the project. Chapter 3 covers the methodology and explains the technical aspects of the work and includes the key formulas and some derivations. Chapter 4 introduces empirical data and critically evaluates the results. Chapter 5 summarizes the findings.

Chapter 2

Literature review

2.1 Efficient Markets Hypothesis

> There can be no ultimate statements in science: there can be no statements in science which can not be tested, and therefore none which cannot in principle be refuted, by falsifying some of the conclusions which can be deduced from them.
>
> *Karl Sigmund Popper*

In general, people like structure and patterns that, sometimes delusionally, they find around themselves. Hence, science relies a great deal on mathematics and physics, which have invented the language for those patterns and laws. But not like the natural world, financial world is governed by peoples' decisions that interact with each other in the tumultuous markets, rather than pure and 'God's given' rules. Nevertheless, also this artificial world, with an imperfect structure and inner chaos, offers a number of puzzles and irregularities to study as well.

When a phenomenon does not have a clear and intuitive reasoning, people tend to, particularly in the old days, seek solutions and answers from a Deity. For finance, and statistics in general, similar case is the Normal distribution. It encapsulates a comfortable world of smooth and continuous density distribution with 'nice' properties.

However, already one of the founding texts of finance (Bachelier 1900, p. 10) notes that "Undoubtedly, the Theory of Probability will never be applicable to the movements of quoted prices and the dynamics of the Stock Exchange will never be an exact science". Nevertheless, the paper lays out the mathematical foundation of arbitrage trading. This suggests, that the origins of modern finance had a little contradicting basis. Bachelier famously states that the "mathematical expectation of a speculator is nil", which spurred the mainstream theory of finance of the previous century, known as the Efficient Market Hypothesis.

Bachelier's work was forgotten for some 60 years, until Leonard J. Savage sent it to MIT's Paul Samuelson and others. Sewell (2011) notes that Savage's postcard, in some sense, started the modern financial theory. For instance, it directly led to the derivation of the famous Black-Scholes option pricing formula, as stated by Davis *et al.* (2011).

The prevailing paradigm of the last century was, that the prices of financial instruments are inherently unpredictable, since they quickly incorporate all available information in the market. This was particularly studied in equities, as they provided most of the then volatility. It was believed, that the markets are efficient and any systematic mispricing, that would lead to a sure profit of a speculator, does not occur. Overall, the theory of efficient markets is rigorously discussed in Malkiel & Fama (1970).

Two main threads of research in EMH are, that prices follow the Random walk process, a concept introduced in Pearson (1905), and that they possess the martingale property. Martingale property states, that the best estimate of a future price is exactly the last seen value, as proposed by Ville (1939) and Samuelson (1965). If unconditionally true, it would discourage any profit–seeking speculation. Random walk as price generating engine in finance is thoroughly inspected by Kendall & Hill (1953).

Financial returns were concretely assumed to follow a Normal distribution. Alexander (1961) finds autocorrelation in monthly returns of cotton on Chicago exchange, that Kendall & Hill (1953) explain with a Random walk process. Osborne (1959) researches logarithmic prices that, in his conclusion, follow a Brownian motion.

For a more detailed perspective and historical connotations of the development of the EMH, see Sewell (2011). Although this notion was nurtured throughout the last century, there was an opposing theory, which did not receive almost any attention.

2.2 Non-Gaussian world

> I shall replace the Gaussian distributions throughout by another family of probability laws, to be referred to as "stable Paretian".
>
> *Benoit B. Mandelbrot*

In a financial world governed by a normally distributed returns, 99.7% of observations fall within 3 standard deviations from the mean. Those boundaries are deterministic and finite depending on the mean and variance parameters. In finance, it has been assumed rather blindly, that asset returns follow the Gaussian laws. This assumption lies, for example, in the core of the well–known Black–Scholes–Merton option pricing formula, where the underlying's returns subscribe to normal density. However, history shows plenty of examples where this assumption is violated.

Although there was a lonely voice in the academic debate claiming, that the financial world is not Gaussian. But Benoit B. Mandelbrot did not manage to turn the attention on his side. The mainstream approach would stay blinded within the three sigmas. To put it into perspective; Larson (1960) fits normal distribution to corn futures returns and finds "excessive number of crashes as fas as 8 or 9 deviations from the mean". That is very alarming, although not uncommon, as the probability of those events is virtually zero under the Normal distribution. The then researchers would mostly disregard it, being adamant that Normal distribution is the right one. They would massively underestimate the probability of crashes – extreme positive returns as well, but nobody would be really complaining by those, particularly in practice. The reasons behind this consistency were both methodological and technical. Some of them are pointed out in Chapter 3.

Key man in the dispute of normality of returns was Benoit Mandelbrot. "We have witnessed crashes in the last century which, according to a classic Gaussian model should not have repeated as they did in the 1980's and 90's. The worst Dow Jones Index performance of the century closed on October 19^{th}, 1987[1] with a plunge of 22.6%, had the probability of occurrence less than 1 in 10^{50}, (Mandelbrot & Hudson 2007, p. 34).

Mandelbrot inspected prices of cotton on the New York stock exchange from 1880 to 1940 without assuming that Gaussian distributions drive the returns. Instead, Mandelbrot suggests that a heavily tailed Cauchy distribution is a better fit and that prices are "almost surely almost everywhere discontinuous" (Mandelbrot 1963b, p. 417).

Moreover, in his lifetime work, he disputes the assumption of constant volatility, generalises Cauchy distribution to Stable distributions and introduces concepts of long–term persistence and self–similarity in returns. Those concepts are building blocks of Fractal Market Hypothesis, an opposing theory to EMH Mandelbrot (1963b;a; 1967).

Nowadays, there exists more evidence to the contrary of EMH, for example Kristoufek & Vosvrda (2013) propose a method to measure efficiency and find 'all' markets rather far from being efficient.

Yet, Mandelbrot's efforts were not to vain. In particular, his research brought more time–series based models into finance; for instance detection of long–memory in the markets, based on a famous paper by Hurst (1951). It shows that a signal can persist for a long time in the series, thus opposing the idea that news are incorporated in the price and then 'vanish'. Long memory is very well summarised for example in Beran (1994).

Research in finance has ever since been adopting methods from physics, particularly signal processing, data mining and recently machine learning. Those go hand in hand with investors looking into automatised strategies for trading and more advanced analysis, such as Wavelets. Those decompose signals on the time and frequency domain simultaneously and enable investors to inspect correlations on different time scales quite easily. For details, see Gençay *et al.* (2001). In portfolio theory, modern methods are being developed in order to get a better estimate of the covariance matrix, such as the Shrinkage estimator, or Random Matrix Theory Edelman & Rao (2005).

[1] Usually referred to as Black Monday

2.3 Portfolio theory

> Since there were two criteria - expected return and risk - the natural approach for an economics student was to imagine the investor selecting a point from the set of Pareto optimal expected return, variance of return combinations, now known as the efficient frontier.
>
> *Harry Markowitz*

Should the world be a deterministic machine, as described almost 400 years ago by Descartes, a life of an investor would be an easy one. In this world, an investor only needs to select to their portfolio assets with the highest returns in the maximum quantities they can. In this world, there would be no benefit in diversification. In this world, diversification does not reduce risk of a portfolio, as described in the revolutionary work by Markowitz (1952).

Markowitz departed from the typical view on the economy, particularly in a microeconomic sense. Instead of inspecting the behaviour of firms and their production, he takes on the perspective of investors. Moreover, his ideas could be used by anyone with access to past stock prices and a computer with a decent linear algebra engine. This was unprecedented in many ways, and was a stepping stone for finance; to be able to become open to the general public one day. More importantly, Markowitz was one of the first to incorporate uncertainty to then's problems at hand Markowitz (1990).

Markowitz is answering the question of what is the optimal allocation of resources to a set of assets, while minimising the risk to a given level of return. In the original setup, as is a common practice today, the risk is approximated as the standard deviation of the assets' returns. The minimisation problem can be transformed into maximising the return to a given level of risk as well.

Quite interestingly, Markowitz shows, that there always exists a Global Minimum Variance portfolio in the mean–variance space, that is independent of the returns. The result, of either of the mentioned problems, is a vector of weights indicating a relative allocation of the funds to assets at hand. This framework has, of course been scrutinised in the academia and built upon.

Couple of alternative portfolio theories were developed by Lee (1977) or Kraus & Litzenberger (1976). Most notably, extension to Sharpe's model that assumes all investors to be optimising in the Markowitz way, is described in Sharpe (1964). Further more, this was elaborated on by Merton (1973) to allow for inter–temporal optimisation. For a more thorough history on the contemporary portfolio theory, see Elton & Gruber (1997).

Although, it is still a common practice, to work in the mean–variance framework, this thesis is looking into different risk measures. In the past two decades, the main risk measure was Value at Risk, particularly popular due to its intuitive interpretation. However, more mathematically sound approach is to use Expected shortfall, which is the main risk measure in this thesis.

2.4 Modern risk management

> The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning. We may construe them in self–serving ways that are detached from their objective reality.
>
> *Nate Silver*

Risk management has focused over the past couple of years very much on capital adequacy. It estimates how much capital an investor, insurer, or particularly a bank, needs in order to remain solvent. In mathematical terms, a common approach is to fit a distribution to the returns and calculate possible losses and their respective probabilities from specified density function.

Introduced in Markowitz (1952), risk has since been commonly approximated by a standard deviation (SD) of the asset's returns. This measure of dispersion around a mean, is quite suitable, should the markets behave very calmly. Unlike Value at Risk (VaR), it has appropriate mathematical properties, as defined in Artzner *et al.* (1999). Unfortunately, it hardly describes anything else beyond how far we need to look in order to see practically all possible outcomes of a distribution. In Gaussian world, it only requires three such steps and anything further is quite beyond our reach. Moreover, they completely disregard probabilities of such events, which are in fact non-zero.

Value at Risk, although a very popular measure, has undesirable mathematical properties. It lacks subadditivity and convexity, as defined by Artzner *et al.* (1999), which will be elaborated upon in Chapter 3. In effect, VaR states, at a certain probability level, what is the maximum loss one can expect. Or rather, what is the amount such that, with some probability (typically 95% or 99%), the loss will not exceed VaR – hence how big a sum is at stake to be lost at most. Despite its critique, VaR was the regulatory standard in finance for most of this century. Now, with the advent of Basel III accord, the industry standards were shifted towards conditional Value at Risk (cVar)[2], also known as Expected Shortfall.

Although, it is possible to optimise a portfolio with respect to VaR, the problem becomes ill–posed, as explain Rockafellar & Uryasev (2000) and Mausser & Rosen (2000). That means, it is not continuous with respect to its initial conditions and resulting weights can, quite often, diverge. Hence, an investor is to sell or buy infinite amount of an instrument, which is not feasible. Yamai *et al.* (2002) also notes that VaR misleads investors when they maximise their expected utility of a portfolio. Examples of infeasibility of VaR are well illustrated in Frey & McNeil (2002).

The crucial book on coherent risk measures by Artzner *et al.* (1999) introduces the smallest and law invariant measure, to replace the VaR, which is Expected Shortfall. Kondor *et al.* (2007) shows that optimising Expected Shortfall, or conditional VaR is almost always well-posed problem. That means that the solution to the problem (which is unique and exists) does not change drastically, if the initial conditions change a little bit. Expected Shortfall gives information about the whole tail of the return distribution, beyond the VaR level, which it hence reports as well.

2.5 Random Matrix Theory

> In this job you really need to know only four things: addition, subtraction, multiplication, and division. And most of the time you can get by without division!
>
> *Emanuel Derman, My Life as a Quant*

The cornerstone of the classical portfolio theory is the measure of co–movement between assets. The covariance, or correlation matrix, is the most

[2]There is nuance in the literature, but for the purpose of this thesis, the terms are interchangeable.

crucial piece in the optimisation problem. Particularly, in case of searching for Global Minimum Variance (GMV) set of weights, which is independent of the expected returns of the assets.

However, anything we say about the covariance matrix of real data, is always just an estimate. Since all financial time series ares just a single realisation of a random process, which is measured only on a truncated time domain, the estimates inherently carry some degree of error. The Random Matrix Theory (RMT) attempts to remove some of the noise and calculate more accurate covariance matrix. The amount of noise in the problem is often in literature proxied as $Q = \frac{T}{N}$, where T is the number of observations of N assets, for example in Pafka & Kondor (2003).

The estimation of covariance matrix is quite a complex problem. It needs to calculate $\frac{N^2}{2}$ parameters from $N \times T$ data points, which is of $\mathcal{O}(N^2T)$ complexity with Schoolbrook algorithm. However, for cases where $N \approx T$, the problem can also become ill–posed and the errors can become enormous. Tracy & Widom (1994) or Pafka & Kondor (2003) show, how particularly in those cases, RMT can yield very good results in removing it.

The complexity of covariance matrix estimation has also been studied extensively. The dimensionality issue is well described in Frankfurter *et al.* (1971) or Dickinson (1974). Common idea is to reduce the complexity problem by clustering. That is to group similar assets together and treat them as one – this unsupervised machine learning method relies in this case on macroeconomic or industrial features as explained by Elton *et al.* (2009). Other methods rely on Principal Component Analyse or Bayesian shrinkage as discussed for instance in Jorion (1986). The empirical part of this thesis considers relatively small scale of the number of assets and its respective history to avoid further computational costs. This work predominantly focuses on removing of the noise in the estimation with Random Matrix Theory methodology.

This method is an extension of particle physics research about random matrices. The first important result in the field was presented by Wigner (1951). Wigner found that eigenvalues of large symmetric matrices, with independent random entries of Gaussian origin, adhere to a semi-circle distribution in the limit of $n \to \infty$, when the entries are appropriately scaled to zero mean and unit variance. Nowadays known as the Wigner's Semicircle Law and was generalised to include distributions with finite second moment.

Marčenko & Pastur (1967) introduce a band for eigenvalues of a covariance matrix of purely random and uncorrelated processes – also known as Wishart matrix, making a transition from Wigner's work to the modern RMT. In other words, should the covariance matrix be estimated on i.i.d processes, the matrix's spectrum is known. Marcenko-Pastur band is derived from a Marcenko density of eigenvalues of the Wishart matrix and is closely linked to the Wigner's Semicircle Law.

Plerou *et al.* (1999), Plerou *et al.* (2002) and Laloux *et al.* (1999) pioneer those ideas to be used in finance and offer first empirical results. Bouchaud & Potters (2003) provides crucial evidence which extends the use of Marcenko-Pastur band beyond empirical covariance matrices assuming Gaussian distributions of the underlying processes. Hence, there exists a closed form formula for determining, whether or not an eigenvalue is attributable to random process, or not, see Laloux *et al.* (2000) or Pafka & Kondor (2003). Then the matrix with trimmed eigenvalues can be reconstructed back to its original space and used in for instance portfolio optimisation, as will be shown in the next chapter.

Another important piece in the RMT puzzle is the very large eigenvalue for empirical phenomena. The sharp edges of the M-P boundary work precisely in the asymptotic case when the number of observations as well as observed objects goes to infinity. Moreover, the law assumes there are no fat-tails in their underlying distributions of the entries, as remarked by Arous & Guionnet (2008). In the empirical cases, the covariance matrices have a number of eigenvalues outside the M-P band, which are assumed to be the signal carriers.

Particularly in finance, there is always one value that is much larger than any other. In our cases, it is often interpreted as the mode of the market, or the market trend. Plerou *et al.* (1999) find that corresponding eigenvector has typically uniform distribution of its elements. That is attributed to equally invested portfolio, because eigenvector can be interpreted as a portfolio allocation of the capital, with its entries as weights of the portfolio. Hence there is a space between the largest eigenvalue and the upper limit of the M-P band, that contains the signal carriers. It is well studied by Tracy & Widom (1994) and assert that the largest eigenvalue follows a famous Tracy-Widom distribution. Important work by Johnstone (2001) binds the RMT and Principal Component Analysis, making RMT particularly appealing to other fields as the relevance of the Tracy-Widom law for the high eigenvalues were confirmed outside of finance.

This research put the foundation stone to the RMT as is used today in finance. It shows how to compare the eigenvalues spectrum of the covariance matrix to one calculated from random noise. It also gives the instructions, although not very clear, how to denoise the typical Pearsonian estimate of the true covariance matrix to a better one.

Ideally, this reconstructed matrix will be a better approximation of the 'true' matrix and yield less risky portfolios, which is one of the main hypotheses of this thesis. Bun *et al.* (2017) stress the necessity of regularisation of the covariance matrix, particularly in higher dimensions in terms of the number of observed series. They elaborate on the argument that in the era of Big data, scientists often face the case of $Q \approx 1$, in which they find the RMT results very helpful in making precise statements about the empirical covariance. The paper also presents the fundamental mathematical derivations from particle physics to the financial markets.

Chapter 3

Methodology

3.1 Creating financial market

The main hypothesis of this work is to inspect whether denoising of the covariance matrix can improve the portfolios created with it. The main obstacle to measuring the covariance is that the data generating process is unobservable. Moreover, we only see one realisation of a random process and on top of that within a finite time interval. In order to validate the covariance measures, it is necessary to create a synthetic world, where the covariance matrix is known and then estimate it. Then it is possible to compare estimates of RMT and Pearson correlation matrices.

Definition 1. *Let $M \in \mathbb{R}^{n \times n}$. We say it is a correlation matrix, if it is symmetric and its quadratic form is positive semidefinite, with a unit diagonal. Also, $\forall i, j, i \neq j$, the (i, j) entry is the Pearson correlation coefficient between two random variables x_i and x_j.*
We say the same matrix is a covariance matrix $\mathbb{C}$, or Σ, if the (i, j) entry is Pearson covariance between random variables x_i and x_j with respective variances on the diagonal.

In order to create correlated sequences of observations, we need to specify the covariance matrix. By definition every correlation matrix has to be symmetric and positive semi-definite. But it is not straightforward to simulate a matrix with those properties. Randomly generated values have very little chance to form a PSD system of vectors, with growing size of the matrix this probability decays very quickly. Note, that we can, and will, use the terms of covariance and correlation almost interchangeably. As correlation is just covariance scaled by the standard deviations of the random variables, as long as we keep those values, we can switch from one to another easily.

Definition 2. *Let $M \in \mathbb{R}^{n \times n}$. Then M is Positive Semi-Definite (PSD) if and only if M is symmetric and*

$$x^T M x \geq 0$$

for all $x \in \mathbb{R}$.

We aim for the correlations to have realistic distribution and the matrix to be a true correlation matrix. Hirschberger *et al.* (2007) asserts that for more than 50 assets, it is virtually impossible to find such a matrix by randomly generating the correlations from whichever distribution and setting the main diagonal to one. Upon constructing such a matrix, it needs to be adjusted to have the necessary properties. Higham (2002) proves the solution to this kind of problem:

For any symmetric $A \in R^{n \times n}$, we compute the distance γ as such:

$$\gamma(A) = \min\{||A - X|| : X \text{ is a correlation matrix}\} \tag{3.1}$$

and find the matrix X that achieves this distance under weighted Frobenius norm $||A||_F^2 = \sum_{i,j} a_{i,j}^2$, respectively:

$$||A||_W = ||W^{1/2} A W^{1/2}||_F \tag{3.2}$$

where W is a symmetric positive definite matrix.

Then the algorithm a finds the solution which is closest to the original matrix whilst being in the intersection of sets of PSD matrices and unit diagonal matrices. Since both are closed and convex, the solution is always unique under the given measure. Higham (2002) solves this problem in a following fashion, wherein the complete proof is presented.

Theorem 1. *The correlation matrix X solves* (3.1) *if and only if*

$$X = A + W^{-1}(VDV^T + diag(\theta_i))W^{-1} \tag{3.3}$$

where W is from Equation 3.2, $V \in \mathbb{R}^{n \times n}$ has orthonormal columns spanning null(X), $D = diag(d_i) \geq 0$ and the θ_i are arbitrary.

We use implementation of this algorithm as the *nearPD* function from *Matrix* package by Bates & Maechler (2016), which can also generate covariance matrices in the same fashion by scaling them to correlation and back.

3.1.1 Simulating financial time series

As we are able to create a valid correlation, or covariance, matrix, we can proceed to generate financial time series with given covariance structure. Despite the critique of EMH and its findings, for this case only, we shall assume that the assets follow well known Geometric Brownian Motion as their price driving process. In order to obtain GBM prices with a certain covariance, their increments are generated from multivariate normal distribution with given Σ.

Quantiles of size $(T \times N)$ are generated from the following density function:

$$f(X) = \frac{1}{\sqrt{2\pi^n \det(\Sigma)}} \exp^{[-\frac{1}{2}(X-\mu)^T \Sigma (X-\mu)]} \sim N(\mu, \Sigma) \tag{3.4}$$

The random draws are calculated with *rmvnorm* function in the *MASS* package for *R* by Venables & Ripley (2002). The choice is motivated by the underlying factorisation of Σ. While most other implementations use Cholesky factorisation, MASS employs spectral decomposition foregoing speed for stability, since the generated covariance matrices can be close to singular, stability of the decomposition is important.

Subsequently, those are used to simulate the Geometric Brownian Motion as stock prices. That is a well–known solution to an Stochastic Differential Equation:

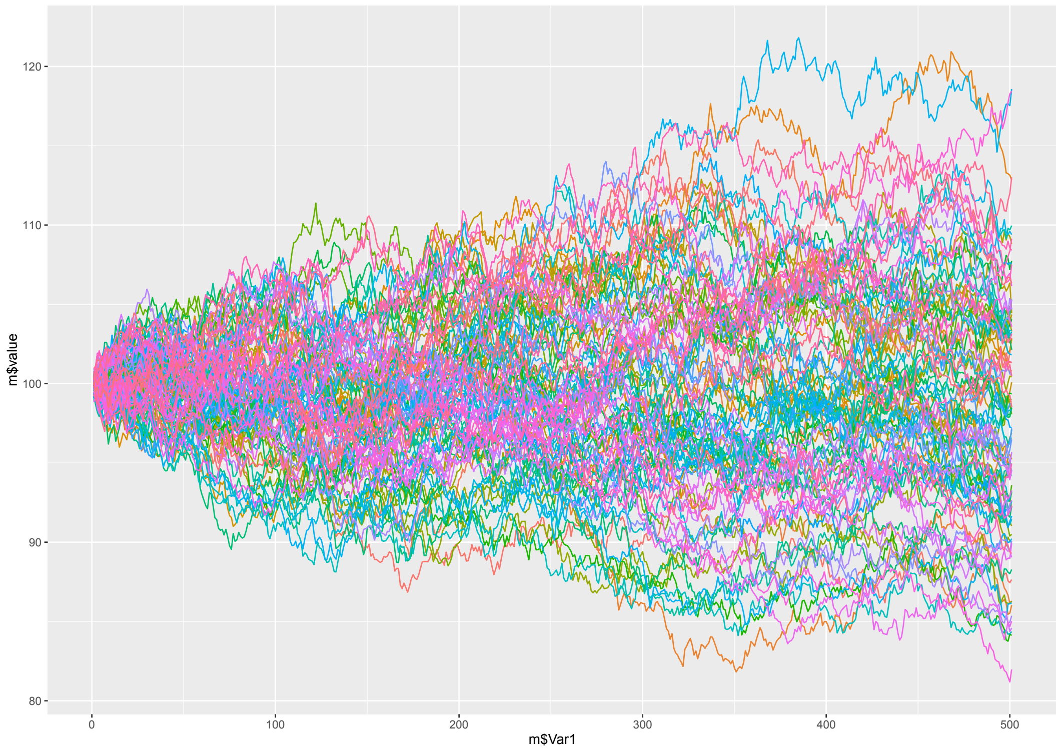
$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.5}$$

where W_t is a Wiener process and μ is the market drift and σ is the volatility of the process with S_t as the price in time t. The covariance structure is imposed in the calculation of the Wiener process which has correlated increments.

Hence, markets of any size in terms of assets or time lengths can be easily generated. This is a crucial step to validate the hypothesis about RMT yielding more precise estimates of the true covariance matrix than Pearsonian estiamtes.

Let us present one simulation of the GMB market to also visually verify its relevance:

Figure 3.1: Simulated GBM market



Therefore, estimation techniques of $\mathbb{C}$ can be validated. Since the true covariance structure of the market is known, we can measure how far off it our estimate is. In order to do that, an appropriate measure of matrix similarity needs to be implemented.

3.1.2 Measuring the covariance matrix

To be able to compare two matrices, or rather distributions, we use the Kullback-Leibler divergence, also known as distance (K-L distance). It is a measure originating in information theory. Intuitively, it measures how much information is lost, if we use Q distribution as an approximation of P distribution, given P as the reference. Since by estimating the covariance matrix, in effect we are estimating the multivariate distribution of assets' returns with a covariance matrix Σ. Moreover, in case of the simulated GBM, the variates are close to being normal with zero mean. Hence, the K-L distance for normal densities can be used.

In this case, P is the density with simulated covariance matrix and Q the density with either the Pearsonian or RMT estimate.

Definition 3. *Let P and Q be probability densities. Then the Kullback-Leibler distance is defined as:*

$$D_{K-L}(P||Q) = \sum_i P(i)log\frac{P(i)}{Q(i)} = \mathbb{E}_P[log\frac{P}{Q}] \tag{3.6}$$

where $\mathbb{E}_P$ denotes mathematical expectation with respected to density P, under the classical probability measure $\mathbb{P}$.

Then the D_{K-L} for normal density is calculated as follows.

Theorem 2. *Let D_{K-L} be defined by Definition 3 and assuming P and Q are multivariate normal covarivance matrices as in 3.4, P with known Σ_P and Q is an estimate of P denoted as $\mathbb{C} = \Sigma_Q$ Then the K-L distance is as follows:*

$$\begin{aligned}
&D_{K-L}(P||Q) = \\
&= \mathbb{E}_P[-\frac{1}{2}\log(2\pi^n \det\Sigma_P) - \frac{1}{2}(x-\mu_P)^T\Sigma_P^{-1}(x-\mu_P) + \frac{1}{2}\log(2\pi^n\det\Sigma_Q) + \\
&+ \frac{1}{2}(x-\mu_Q)^T\Sigma_Q^{-1}(x-\mu_Q)] \\
&= \frac{1}{2}(\log(\frac{\det\Sigma_Q}{\det\Sigma_P}) - (\Sigma_P^{-1}\Sigma_P) + (\Sigma_P^{-1}\Sigma_P) + (\mu_P^T\Sigma_Q^{-1}\mu_P - 2\mu_P^T\Sigma_Q^{-1}\mu_Q + \mu_Q^T\Sigma_Q^{-1}\mu_Q) \\
&= \frac{1}{2}(\log(\frac{det\Sigma_Q}{det\Sigma_P}) - n + (\Sigma_Q^{-1}\Sigma_P) + (\mu_Q-\mu_P)^T\Sigma_Q^{-1}(\mu_Q-\mu_P))
\end{aligned}$$

The derivation uses properties of matrix trace and elementary properties of the mathematical expectation. For example Duchi (2007) presented the proof in full length.

This subsection demonstrated a framework for validating the use of RMT. We expect that, particularly for noisy matrices, such as when $Q \sim 1$, RMT will perform better than Pearsonian covariance. Throughout this work, $Q = \frac{T}{N}$ with T being the number of observations and N the number of assets, is used to calculate the covariance matrix, or optimise a portfolio.

3.2 Portfolio selection

3.2.1 Markowitz solution

Let us briefly remind ourselves of the usual portfolio solution, that takes as risk standard deviation of its components, so that we can compare it to our optimisation problem of minimising the Expected shortfall.

The Markowitz's solution seeks to find the best allocation decision within the risk–return space which either minimises the risk or maximises the return. This yields a so called Efficient frontier of optimal portfolios. Each frontier has a special case with minimum variance called Global Minimum Variance portfolio – it is also a portfolio with minimum expected return on the Efficient boundary. This portfolio is independent of assets' return and will be always the target in this work. Hence, as Merton (1980) points out, by focusing on this particular portfolio, uncertainty of the risk returns' is removed from the problem. GMV is found using only an inverse of the Pearsonian covariance matrix is the input to the optimisation.

Definition 4. *Let $M \in \mathbb{R}^{N \times T}$ be a matrix of returns of N assets with T balanced observations. The Pearsonian estimate of the covariance matrix, for centred data, is calculated as:*

$$\mathbb{C}^{(N \times N)} = \frac{1}{T} M M^T \tag{3.7}$$

Then the portfolio problem is formulated as:

Theorem 3. *Let $\mathbf{x} = (x_1, x_2, ..., x_n)^T$ be the vector of real-random variables representing returns of risky assets, $\Sigma \in \mathbb{R}^{n \times n}$ be the Pearson covariance matrix of $\mathbf{x}$ and $\mathbf{w} = (w_1, w_2, ..., w_n)^T$ be the vector of optimal weights. The problem of finding GMV portfolio can be concisely expressed as:*

$$\min_{\mathbf{w}} \sigma^2 = \mathbf{w}^T \Sigma \ \mathbf{w} \ s.t. \ \mathbf{w}^T \mathbf{1} = 1. \tag{3.8}$$

The first order conditions from Lagrange optimisation state that:

$$\mathbf{0} = \frac{\partial L(\mathbf{w}, \lambda)}{\partial \mathbf{w}} = 2\Sigma \mathbf{w} + \lambda \cdot \mathbf{1} \tag{3.9}$$

$$\mathbf{0} = \frac{\partial L(\mathbf{w}, \lambda)}{\partial \lambda} = \mathbf{w}^T \mathbf{1} - 1 \tag{3.10}$$

Using 3.9 we solve for $\mathbf{w}$*:*

$$\mathbf{w} = -\frac{1}{2}\lambda\Sigma^{-1}\mathbf{1}$$

and then after multiplying both sides by $\mathbf{1}^T$ *from the left, we can use 3.10:*

$$1 = \mathbf{1}^T\mathbf{w} = -\frac{1}{2}\lambda\mathbf{1}^T\Sigma^{-1}\mathbf{1}$$
$$\implies \lambda = -2\frac{1}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}$$

Finally, we substitute the value of λ *back into 3.9 to find the solution:*

$$\mathbf{w} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} \tag{3.11}$$

This demonstrates that the covariance matrix is indeed the cornerstone of the portfolio selection in the classical sense. Moreover, the GMV solution is very straightforward. In fact, the resulting vector of optimal weights is only a vector of sums of rows. Those sums are scaled by the sum of all elements of the inverse of the covariance matrix.

3.2.2 Minimising Expected shortfall

Since the risk metric used for this thesis is the Expected shortfall, one of the hypotheses is whether a lower risk portfolios can be obtained by directly minimising the ES. This approach is also quite recent and is growing in importance as ES is being integrated in the modern risk management. And unlike the Markowitz's solution, it does not require explicitly to measure the covariance of assets. Moreover, Rockafellar & Uryasev (2002) point out, that minimising the ES is much more suitable, unless the underlying distributions are very simple.

Following Rockafellar & Uryasev (2000), we introduce the loss function of a portfolio $f(\mathbf{w}, \mathbf{y}), f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, where $\mathbf{w} \in \mathbb{R}^n$ is again the vector of selected weights and $\mathbf{y} \in \mathbb{R}^m$ represents the market returns with density $p(\mathbf{y})$. As the authors show, the density do not need to be specified analytically and random sampling is sufficient. Also we introduce the probability of loss not exceeding a threshold α as

$$\Psi(\mathbf{x}, \alpha) = \int_{f(\mathbf{x},\mathbf{w}) \leq \alpha} p(\mathbf{y}) d\mathbf{y} \tag{3.12}$$

The function $\Psi(\mathbf{x}, \alpha)$ is a cumulative distribution function of loss and has desirable mathematical properties. It completely determines the level of VaR as well as cVaR. In general, $\Psi(\mathbf{x}, \alpha)$ is a non-decreasing with respect to α which is also right-continuous. We shall assume that no jumps from the left occur, hence it is everywhere smooth and continuous.

Denoting the level of VaR, which will be introduced in Section 3.5, as ϵ and the level of confidence as α, the key to the optimisation problem is function F_α on $X \times \mathbb{R}$.

$$F_\alpha(\mathbf{x}, \epsilon) = \epsilon + (1 - \alpha)^{-1} \int_{y \in R^m} [f(\mathbf{x}, \mathbf{y}) - \epsilon]^+ p(\mathbf{y}) d\mathbf{y} \tag{3.13}$$

where $[t]^+ = t$ when $t > 0$ but $[t]^+ = 0$ when $t \leq 0$. Moreover, this function is shown to be strictly convex in Rockafellar (2015) and Shor (2012), which ensures a minimum exists and is a global minimum. The the minimisation problem is constructed with the following theorem by Rockafellar & Uryasev (2000).

Theorem 4. *Since $F_\alpha, (\mathbf{x}, \epsilon)$ is convex and continuously differentiable, the $cVaR_\epsilon(X)$ of the loss associated with any $x \in X$ can be calculated from*

$$\phi_\alpha(\mathbf{x}) = \underset{\epsilon \in R}{min}\ F_\alpha(\mathbf{x}, \epsilon) \tag{3.14}$$

Hence, minimising ES for convex and differentiable functions is feasible to minimise numerically. For proof of Theorem 4 and more detailed discussion, see the crucial paper by Rockafellar & Uryasev (2000) or Rockafellar & Uryasev (2002) and Rockafellar (2015) .

3.3 Random Matrix Theory

The usual method for estimating the covariance, or correlation, is the Pearsonian method, se defined in Definition 4. As was noted above, this approach inherently involves a certain level of noise. Laloux *et al.* (2000) shows, that the least risky portfolios, that investors are after, are attributed to the smallest eigenvalues of the covariance matrix. But, those are exactly the ones that fall often within the Marcenko-Pastur band, which means they are in impacted the most by noise in the estimation.

The spectrum in case of empirical data, which do contain some signal, also contains a single large eigenvalue. It can be greater than the second largest value even by an order of magnitude. In the literature, this eigenvalue is often interpreted as the market trend, or market mode. The elements of the corresponding eigenvector, which can be interpreted similarly to portfolio weights, are usually close to being uniformly distributed. Hence they would represent $1/N$ portfolio.

Random matrices and theories around them were pioneered in the fifties in particle physics. Hence, quite a few results are already known and are of a genuine interest in finance, which however took interest in it only recently. The main findings are, the works by Bouchaud & Potters (2003), Laloux *et al.* (1999) and Potters *et al.* (2005).

Elements of the correlation matrix are calculated from a series of length T. "If T is not very large compared to N, one should expect that the determination of the covariances is noisy, and therefore that the empirical correlation matrix is to a large extent random."(Laloux *et al.* 2000, p.1).

RMT compares the spectrum of eigenvalues of a matrix consisted of random $i.i.d.$ covariances to an empirical one. Hence, under the null hypothesis, we can determine whether the covariances measure originate from signal, or just noise. Random Matrix Theory asserts the density of eigenvalues of a Wishart matrix, which subscribes to Marcenko-Pastur distribution.

Definition 5. *We shall denote $\rho(\lambda)$ the density of eigenvalues of $\mathbb{C}$, defined as:*

$$\rho(\lambda) = \frac{1}{N}\frac{dn(\lambda)}{d\lambda} \tag{3.15}$$

where $n(\lambda)$ is the number of eigenvalues of $\mathbb{C}$ less than λ

Edelman & Rao (2005) shows, that for random matrices, so-called Wishart matrices, the density $\rho(\lambda)$ is self-averaging. Also, it is known precisely in the limit as $N \to \infty$ and $T \to \infty$ and for a fixed $Q = T/N \geq 1$:

$$\rho_{\mathbb{C}}(\lambda) = \frac{Q}{2\pi\sigma^2}\frac{\sqrt{(\lambda_{max}-\lambda)(\lambda-\lambda_{min})}}{\lambda} \tag{3.16}$$

$$\lambda_{min}^{max} = \sigma^2(1 + 1/Q \pm 2\sqrt{1/Q}) \tag{3.17}$$

for $\forall\lambda \in [\lambda_{min}, \lambda_{max}]$, where the interval for λ is known as the Marcenko-Pastur band. σ^2 is the variance of elements of M – in our case normalised to unit variance. The scaling is for better numerical properties, as the simulated GMB paths exhibit quite low variances.

The method separates the eigenvalues by the Marcenko-Pastur band, where the larger ones are treated as the signal and those within as noise. In the limit $N \to \infty$ there are no eigenvalues below the edge, however, for finite N, there is a probability of finding few below the edge as well. In our case those values are treated as noisy as well.

The actual process of denoising is not unambiguous. Tola *et al.* (2008) for instance states that to obtain the denoised matrix, the matrix needs to have eigenvalues within the M-P band set to 0 and then "transformed into the original bases". However, there are no precise steps of instructions outlined. Bouchaud & Potters (2009) also notes that the noisy eigenvalues should be set to zero, with no further details. For that reason the calculations were discussed with Prof. Brian Rowe [1]

Theorem 5. *Let us consider $\mathbb{C}$ the estimated covariance matrix and $\mathbb{L}$ the matrix of respective eigenvectors. Let us define $\mathbb{D}$ as the diagonal comprising of the filtered eigenvalues. The ones below the λ_{max} are replaced by their respective mean. To transform the $\mathbb{D}$ into basis of $\mathbb{C}$, we calculate :*

$$\hat{\mathbb{C}}' = \mathbb{L}\mathbb{D}\mathbb{L}^T \tag{3.18}$$

To ensure the resulting matrix is a correlation matrix, let us apply a correction:

$$\hat{\mathbb{C}} = \frac{\hat{\mathbb{C}}'}{\sqrt{(\mathbb{D}')(\mathbb{D}')^T}}$$

[1] Professor of mathematics at CUNY School of Professional Studies and author of RMT library Rowe (2016); r@zatonovo.com

where $\mathbb{D}'$ is the outer product of $\mathbb{D}$ and the unit vector.

3.4 Stable distribution

This thesis replaces the usual assumption about normality of returns with a more general family of distributions. So called α-stable Lévy distributions are far more suitable for modelling financial returns for many reasons outlined in Chapter Chapter 2. Nevertheless, plenty of models and concepts embrace it; from Black-Scholes option pricing engine to RiskMetrics's variance-covariance method to calculate Value at Risk. Stable distributions are rarely used also because there is a bit of confusion in the methodology, as Nolan (2003) points out. This section will attempt to clarify them and will therefore go in a little more detail.

Stable laws where introduce by Lévy (1925) and are also known as Pareto-Lévy distributions, particularly due to Benoit B. Mandelbrot. The most general form is often referred to as α-stable Lévy distribution. Normality of returns is an attractive assumption because of Central Limit Theorem. Gnedenko & Kolmogorov (1954) generalises the CLT by showing that "the only possible non-trivial limit of properly normalised sums of $i.i.d$ increments is the stable law. This result allows the CLT to include increments with infinite variance. Nolan (2012) then defines stable distributions in the following manner:

Definition 6. *A random variable X is stable, or stable in the broad sense, if for X_1 and X_2 independent copies of X, and any $a, b, c \in \mathbb{R} > 0$ and some $d \in \mathbb{R}$, holds*

$$aX_1 + bX_2 \stackrel{d}{=} cX + d \tag{3.19}$$

The random variable is strictly stable or stable in the narrow sense if 3.19 holds with d=0 for all $a, b, c \in \mathbb{R}$. A random variable is symmetric stable if it is stable and symmetrically distributed around 0, e.g. $X \stackrel{d}{=} -X$.

Theorem 6. *A non-degenerate random variable Z is α-stable for some $0 < \alpha \leq 2$ if and only if there is an independent identically distributed sequence of random variables X_1,X_2,X_3,... and constants $a_n > 0, b_n \in \mathbb{R}$ with*

$$a_n(X_1 + X_2 + X_3 + ... + X_n) - b_n \xrightarrow{d} Z \tag{3.20}$$

One of the biggest complication when working with stable distributions is the number of parametrizations in the literature. Nolan (2003) states there are about 6 main parametrizations. All of them are covered in detail in Nolan (1997) and Borak *et al.* (2005). In this work, we restrict ourselves to use parametrisation by Samoradnitsky & Taqqu (1994) because according to Nolan (2001) it has better numerical properties. Also, it is used by Stoyanov *et al.* (2006) to calculate the Expected shortfall.

Definition 7. *$X \sim \mathbb{S}(\alpha, \beta, \gamma, \delta)$, if the log-characteristic function*[2] *of X is given by:*

$$log(\psi(t)) = log(\mathbb{E}(e^{ixt})) = \begin{cases} i\delta t - |\gamma t|^{\alpha} \cdot [1 - i\beta \cdot sign(t) tan(\frac{\pi\alpha}{2})] & \text{if } \alpha \neq 1 \\ i\delta t - |\gamma t| \cdot [1 + i\beta \cdot sign(t) \frac{2}{\pi} log|t|] & \text{if } \alpha = 1 \end{cases} \tag{3.21}$$

where $\alpha \in (0, 2]$ is the index of stability, $\beta \in [-1, 1]$ is the skewness parameter, scale is denoted by $\gamma \in \mathbb{R}, \gamma > 0$, and location parameter $\delta \in \mathbb{R}$.

The parameter α controls the heaviness of tails, or rather the variance of the distribution (rather than the usual σ). The variance of the distribution is always infinite, except for a special case $\alpha = 2$, when converges to Normal distribution $N(\delta, 2\gamma^2)$. This is one source of mistakes even in a renowned literature, which sometimes mistakenly uses $N(\delta, \gamma)$.

The location parameter is denoted δ instead of the usual μ in Normal case, since for $\alpha \leq 1$, the Stable distributions do not have a mean in the usual sense. However, we restrict ourselves to cases when $\alpha \gg 1$ since the empirical data have the α parameter well over 1.5.

Zolotarev (1986) introduces an important parametrisation, so called M, which is jointly continuous in all parameters. Samoradnitsky's version is discontinuous at $\alpha = 1$. The M parametrisation will be useful to showing how to calculate density of a Stable distribution.

3.4.1 Estimation of parameters

In the Gaussian world, there are only two parameters to estimate in order to fit the distribution. In this case, there are four and they are not defined that well. Because stable distributions are only defined via their respective characteristic

[2] Or its equivalent without the logarithm.

functions, they lack a closed form of the density function.

McCulloch (1986) presents a straightforward method of calculating the parameters. Since the parameter space is continuous, the method interpolates between pre–calculated quantiles and infers the data generating specification. This process is very fast, but not very accurate. Since this work assumes data come from a certain distribution, the optimal solution would be to use Maximum Likelihood Estimation. Nolan (2001) finds, as would be expected, that MLE is always more accurate, than the quantile method.

Maximum Likelihood Estimates

The principle of MLE is to determine what is the most probable set of parameters which generated the data at hand. It requires to derive a likelihood function, which depends on the parameters and for the input data and can be maximised on the parameter space.

Definition 8. *Let* $X \sim \mathbb{S}(\alpha, \beta, \gamma, \delta)$. *Let us denote parameter vector* $\theta = (\alpha, \beta, \gamma, \delta)$, *the density function as* $f(x|\theta)$ *and parameter space as* $\Theta = (0, 2] \times [-1, 1] \times (0, \infty) \times (-\infty, \infty)$. *Then the log likelihood function of identically independently distributed stable sample* $X_1, X_2, ..., X_n$ *is given by*

$$l(\theta) = \sum_{i=1}^{n} \log f(X_i|\theta)$$

Nolan (1997) presents a semi closed form of the density function for Zolotarev's (M) parametrisation, which is also used in the calculation of the Expected shortfall shown later. However, its implementation in R by Wuertz *et al.* (2016) is very slow and makes the MLE optimisation unfeasible. Since it needs to be evaluated plenty of times, a faster solution is required. Nevertheless, we can use duality of characteristic functions and density functions in Fourier transform to calculate the density. As the initial starting point for MLE McCulloh's quantile method is used with implementation in *fBasics* package by Wuertz *et al.* (2014).

The author wrote a library which uses the FFT and runs two orders of magnitude faster than in *fBasics*. It uses integration of C code in R via the *Rcpp* package by Eddelbuettel & Francois (2011). Also, Rachev & Mittnik (2000) find that using the FFT approach is much faster if the number of inputs is over

100, or $\alpha > 1.5$, which is the case of this work.

Lambert & Lindsey (1999) propose a modification of the logarithmic characteristic function in equation 3.21 for better numerical properties in the calculation.

$$log(\psi(t)) = i\delta t - \gamma'^{\alpha}|t|^{\alpha} \exp\{-i\beta' \frac{\pi}{2}\eta_{\alpha}\text{sign}(t)\} \tag{3.22}$$

$$\eta_{\alpha} = \min(\alpha, 2-\alpha) = 1 - |1-\alpha| \tag{3.23}$$

where parameters are transformed from 3.21as follows:

$$\beta' = \frac{2}{\pi\eta_{\alpha}} \cos^{-1}\left\{\frac{\cos(\pi\alpha/2)}{\Delta}\right\}$$
$$\gamma' = \{\frac{\Delta\gamma}{\cos(\pi\alpha/2)}\}^{1\alpha}$$

where

$$\Delta^2 = \cos^2(\pi\alpha/2) + \beta^2\sin^2(\pi\alpha/2)$$
$$\text{sign}(\Delta) = \text{sign}(1-\alpha)$$
$$\text{sign}(\beta') = \text{sign}(\beta)$$

The tail and location parameters are unchanged. By applying the Fourier transform to characteristic function in equation 3.22 we obtain the density function as follows:

$$\frac{1}{\pi}\int_0^{\infty} \cos[(\delta - y)\frac{s}{\gamma'} + s^{\alpha}\sin(\eta'_{\alpha,\beta'})]\exp[-s^{\alpha}cos(\eta'_{\alpha,\beta'})]\frac{ds}{\gamma'} \tag{3.24}$$

This integral can be calculated numerically with any quadrature, author's library uses Romberg integration which is also written in C.

3.5 Expected shortfall

As stated previously, ES is becoming the new industry standard as a risk measure. It is replacing Value at Risk which has undesirable mathematical properties. Particularly it does not satisfy coherency as defined by Artzner *et al.* (1999).

Definition 9. *Let us consider a set V of real-valued random variables. A function $\rho : V \to \mathbb{R}$ is a risk measure if it is:*

(i) monotonous: $\forall X, Y \in V, Y \geq X \implies \rho(Y) \leq \rho(X)$

(ii) sub-additive: $\forall X, Y, X + Y \in V \implies \rho(X + Y) \leq \rho(X) + \rho(Y)$

(iii) positively homogenous: $X \in V, h > 0, hX \in V \implies \rho(hX) = h\rho(X)$

(iv) translation invariant: $X \in V, a \in \mathbb{R} \implies \rho(X + a) = \rho(X) - a$

Risk measure is a kind of spectral measure functions which is in most general form defined by Dowd (2007).

$$M_\phi = \int_0^1 \phi(p) q_p dp \tag{3.25}$$

where $\phi(p)$ is the weighting function. It is also referred to as the risk spectrum, or risk-aversion function.

Expected shortfall as well as Value at Risk are spectral measures, where VaR places all weight on just one value, while ES takes an average of values below a certain threshold. Acerbi (2002) introduces restrictions on $\phi(p)$ so that M_ϕ is coherent;

- *Non-negativity:* $\forall p \in [0, 1], \phi(p) \geq 0$
- *Normalization:* $\int_0^1 \phi(p) dp = 1$
- *Weakly increasing:* Weights do not decrease as the quantiles increase

To define ES as a spectral risk measure, M_ϕ, we set $\phi(p)$ to the form:

$$\phi(p) = \begin{cases} \frac{1}{\epsilon} & \text{if } p \leq \epsilon \\ 0 & \text{if } p > \epsilon \end{cases} \tag{3.26}$$

where ϵ is the level of significance. This function certainly satisfied Equation 3.26. Then with Equation 3.25 we derive the final expression for Expected shortfall which is a coherent measure by Definition 9. It is worth noting that in this thesis we consider losses to be negative returns, rather than positive values. Hence the definition considers values less than a negative VaR, which is, however, being treated as a positive quantity in the usual sense.

Definition 10. *Let X be a financial time series*

$$ES(X) = \frac{1}{\epsilon}\int_0^{\epsilon} VaR_{\epsilon}(X)d\epsilon \tag{3.27}$$

where $VaR_{\epsilon}(X)$ is a value satisfying $\mathbb{P}(X \leq -VaR_{\epsilon}(X)) = \epsilon$

3.5.1 Expected shortfall for stable distribution

In the last past we only need to introduce the formula to calculate the Expected shortfall for the distributions of choice, which is instrumental to this thesis. The whole optimisation problem is well covered by Stoyanov *et al.* (2006), where detailed proof for the following theorem can be found. Also, let us note that Stoyanov denotes Expected shortfall $cVaR_{\epsilon}(X)$ rather than $ES_{\epsilon}(X)$, so we follow his notation in this case.

Theorem 7. *Let $X \sim \mathbb{S}(\alpha, \beta, 1, 0)$ with $\alpha > 1$. If $VaR_{\epsilon}(X) \neq 0$, then the cVar of X at significance level ϵ admits the following integral form:*

$$cVaR_{\epsilon}(X) = \frac{\alpha}{1-\alpha}\frac{|VaR_{\epsilon}(X)|}{\pi\epsilon}\int_{-\bar{\theta}_0}^{\pi/2} g(\theta)\exp(-|VaR_{\epsilon}(X)|^{\frac{\alpha}{\alpha-1}}v(\theta))d\theta \tag{3.28}$$

where

$$\begin{aligned}
g(\theta) &= \frac{\sin[\alpha(\bar{\theta}_0+\theta)-2\theta]}{\sin[\alpha(\bar{\theta}_0+\theta)]} - \frac{\alpha\cos^2(\theta)}{\sin^2[\alpha(\bar{\theta}_0+\theta)]},\\
v(\theta) &= (\cos[\alpha\bar{\theta}_0])^{\frac{1}{\alpha-1}}\left(\frac{\cos\theta}{\sin[\alpha(\bar{\theta}_0+\theta)]}\right)^{\frac{\alpha}{\alpha-1}}\frac{\cos[\alpha\bar{\theta}_0+(\alpha-1)\theta]}{\cos(\theta)},\\
\bar{\theta}_0 &= \frac{1}{\alpha}\arctan(\bar{\beta}\tan\frac{\pi\alpha}{2}), \textit{ and}\\
\bar{\beta} &= -sign(VaR_{\epsilon}(X))\beta
\end{aligned}$$

Furthermore, if $VaR_{\epsilon}(X)$ =0, then

$$cVaR_{\epsilon}(X) = \frac{2\Gamma(\frac{\alpha-1}{\alpha})}{\pi-2\bar{\theta}_0}\frac{\cos\bar{\theta}_0}{(\cos[\alpha\bar{\theta}_0])^{1/\alpha}} \tag{3.29}$$

For symmetric cases with $\beta = 0$, the equation can be broken down to:

Theorem 8. *If* $X \sim \mathbb{S}(\alpha, 0, 1, 0)$ *with* $\alpha > 1$ *and* $VaR_\epsilon(X) \neq 0$, *then* $cVaR_\epsilon(X)$ *admits the following representation:*

$$cVaR_\epsilon(X) = \frac{\alpha}{1-\alpha} \frac{|VaR_\epsilon(X)|}{\pi\epsilon} \int_0^{\pi/2} g(\theta) \exp(-|VaR_\epsilon(X)|^{\frac{\alpha}{\alpha-1}} v(\theta)) d\theta \tag{3.30}$$

where

$$g(\theta) = \frac{\sin[(\alpha-2)\theta]}{\sin[\alpha\theta]} - \frac{\alpha \cos^2(\theta)}{\sin^2[\alpha\theta)]},$$
$$v(\theta) = \left(\frac{\cos\theta}{\sin[\alpha+\theta]}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos[(\alpha-1)\theta]}{\cos(\theta)},$$

If also $VaR_\epsilon(X) = 0$, *then*

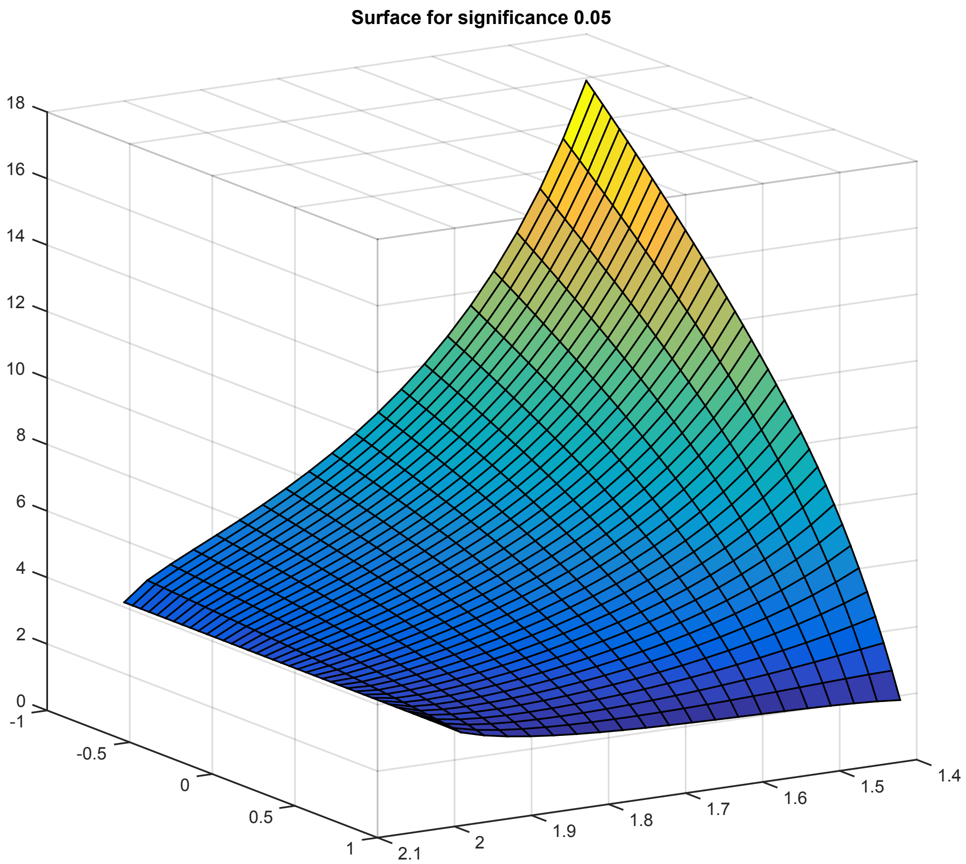
$$cVaR_\epsilon(X) = \frac{2\Gamma(\frac{\alpha-1}{\alpha})}{\pi}$$

The proof is lengthy, yet relatively simple. It uses Nolan's integral representation of the density described in Equation 3.22 with appropriate reparametrization to Samoradnitzsky's version and coherency properties.

Numerical obstacles

To validate the calculations, we replicated the ES surfaces from Stoyanov's paper. The integral has singularity on the left side of the interval and is little complicated to evaluate using some quadratures. In R, the first choice for integration method is the built-in function *integrate* that uses a mix of adaptive quadratures. However, it does not always converge and in some cases it takes up extreme integration error, which was difficult to reduce. Other integrating functions were tested with little or no benefit to the problem.

Figure 3.2: Replication of original surface



The left axis is the β parameter and the other is the tail heaviness parameter α. The surface was computed in R and only plotted in MATLAB for its convenient plotting functionality. The surface does not decrease as the α parameter decreases, since it is creating fatter tails, the ES cannot decrease. The only exception are values of β close to 1, where the singularity causes underestimation of the Expected shortfall. Those however do not occur very often when working with empirical data. Nonetheless, the integration fails most of the time in situations when $\alpha < 1.5$ and $\beta \notin [0, 0.5]$, which does not happen often with empirical data.

Upon consulting with Dr Stoyanov by email, he pointed out that his team used MATLAB's *quandl* function, which uses adaptive Lobatto quadrature method. Indeed in MATLAB the whole surface can be replicated well and the integral fails to calculate only in a small number of cases. R supplies the same quadrature, but it never finishes the computation. Hence, checks were written that should the attempt to find a solution for the closest set of parameters, sometimes just truncating the parameters' values helps a great deal.

If that does not reduce the absolute error under 10^{-2}, we simulate quantiles of the specified distribution within the ES tail and average them. A sequence of 500 probabilities from the interval $(0, \epsilon]$ is taken and respective quantiles calculated. For that, *stabledist* by Wuertz *et al.* (2016) package is used.

When compared to Stoyanov's results calculated in MATLAB then our results admit MSE of 0.00268, with SD 0.0033 and median error 0.00123. Those are worked out as differences in 1000 data points randomly selected in the '$\alpha - \beta$ space.

3.5.2 Non-parametric approach

We introduced a calculation framework which assumes that financial returns follow a certain distribution. The results then show, what are the ES values, should we observe the time series for an infinitely long time, so that each value of the real axis is observed. That is of course impossible. Therefore we take value from Stoyanov's algorithm as the ground truth and calculate also realised risk, which an investor faces while holding a portfolio for a limited amount of time.

Definition 11. *Let $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ denote the order statistics in ascending order corresponding to financial returns, $X_1, X_2, ..., X_n$. The realised Expected Shortfall as follows:*

$$\widehat{cVaR}_\epsilon(X) = (\sum_{i=1}^{[n\epsilon]} X_{(i)})/(n - [n\epsilon]) \tag{3.31}$$

where [y] is the floor operator.

Non-parametric methods are suitable because they allow us to avoid the pitfalls of wrongly specifying the distribution. When correct, the distribution parameters is the only information necessary to completely describe the phenomenon. We are able to evaluate any scenario, or precisely predict losses. However, if mistaken, all the inference made contains at least considerable errors.

Also, non-parametric approach is far more intuitive and straightforward to calculate – it only requires computing order statistics. It also accommodates to any shape the data might have. In our case, the calculation of Expected shortfall sometimes fails if the series is too short and resembles a multimodal distribution, because the MLE parameters are bordering on the feasibility set. Yet, the input date are also its severe limitation since should we for instance had observed a quiet period, the resulting risk metrics would be significantly underestimated, because unobserved values would have zero 'probability'. Whereas in the parametric case, even unobserved events are assigned a probability and therefore taken into account.

Chapter 4

Results

4.1 Data

The dataset was selected to represent a market with a period of high volatility and long enough history to be able to create a medium size portfolio. Small portfolios, or rather portfolios with low number of represented sectors are unsuitable for they limit the amount of risk that could be diversified. Hence, all constituents of the SP 500 index were considered and then only the stocks with full history between January 2007 and December 2011. The time period was chosen purposefully to include a period of high volatility in order to measure how each method would cope with sudden drop in market prices and then the subsequent recover. Out of the full sample, a random subsample of 60 stocks were selected, in order allow for large Q parameter, where $60 \times Q$ observations of prices are needed.

The data were downloaded from Yahoo! finance database. In order to account for stocks' splits and dividends, we use Adjusted daily close prices, provided by Yahoo!, and always consider logarithmic returns for they are more suitable from the computational point of view, than . The summary of the stocks' individual characteristics are in the Appendix as Table 1 and Table 2.

The portfolio selection places no restrictions on short selling of the assets, so that the optimisations remain linear and solvable with standard tools. For this work is concerned primarily with managing large losses, transactions costs are disregarded. Therefore, the strategies presented are not suitable in practice for individual investors, particularly for the availability and arbitrary duration of short positions.

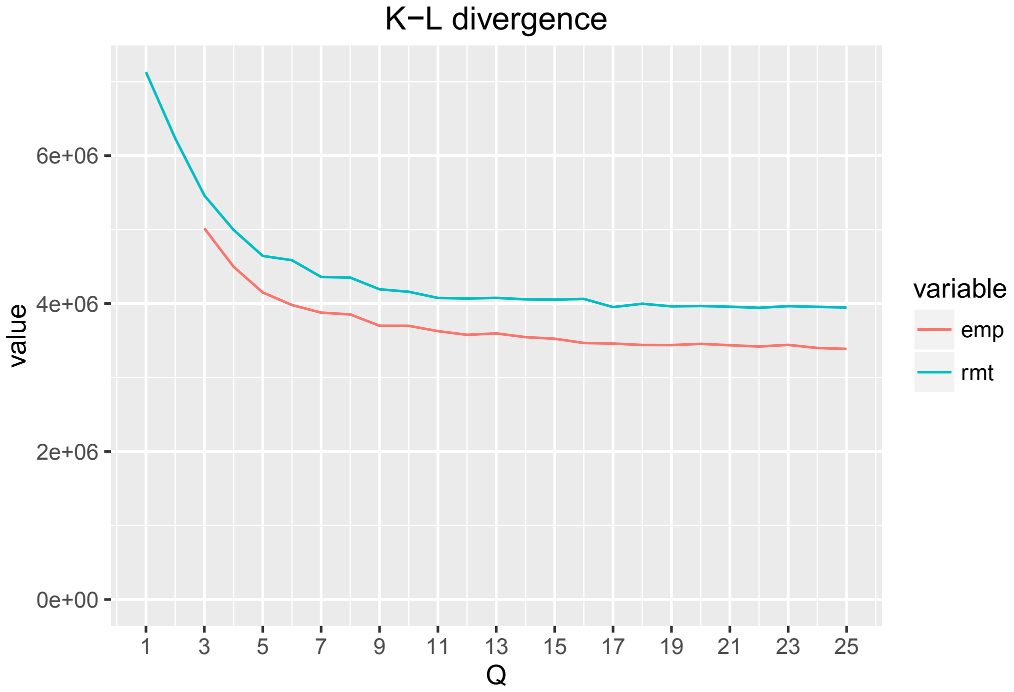
4.2 Relevance of Random Matrix Theory

Denoising of the covariance matrix for use in finance is a relatively new and therefore only limited number of empirical results are available. One of the hypothesis of this work is to inspect actual relevance of RMT to the problem. Let us assume that the covariance structure of the market is known. In other words, the covariance matrix is observed and therefore, degree of error of the estimates can be measured.

As outlined in the previous chapter, we simulated a valid covariance matrix from which a set of asset prices were generated using the Geometric Brownian Motion. Using those, the covariance matrix is estimated and compared to the true matrix in its initial and denoised form. The distance measure is the Kullback-Leibler distance. The distances from the true matrix are measured with varying Q as the noise coefficient, and are to converge to zero eventually. The Q is varied from 1 to 100 on a simulated set of 100 time series.

4.2.1 Simulated data

Figure 4.1: K-L distances for simulated financial market



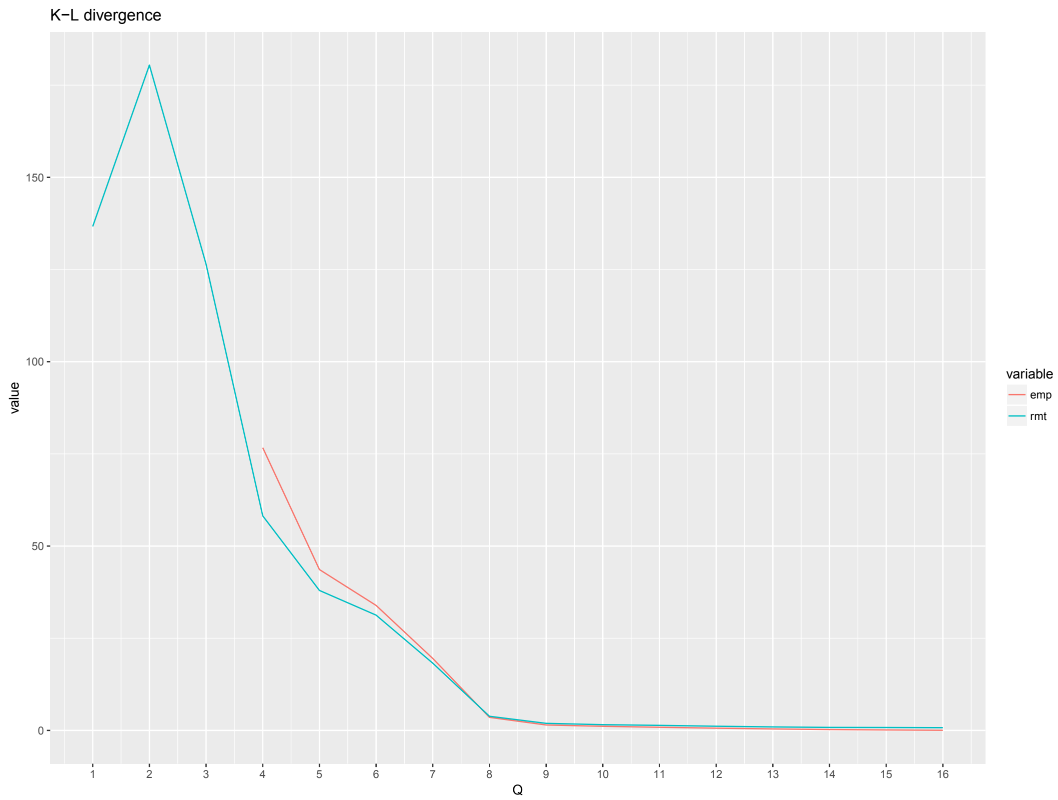
Note that for $Q \approx 1$ both the values are off the scales, with the empirical covariance diverging to infinity, while the RMT has a finite value, although a very high one. The empirical covariance drops to scale only for $Q = 3$. In this case, the values obtained are very large – particularly compared to empirical data – due to the numerical constraints imposed by the algorithm which creates the original matrix. The matrices generated by Higham's have very small determinants and are close to singular. This causes the term $\log \frac{\det \Sigma_Q}{\det \Sigma_P}$ in D_{K-L} to attain large values and not reach 0.

When compared to empirical series, the simulated returns follow somewhat thinner distribution around zero, which is caused by the underlying Brownian Motion with relatively low variance and quite granular discretisation of the stochastic process. However, since the $K - L$ distance is a dimensionless measure, the scale does not really matter. More importantly, the simulated data show that RMT method outperforms the empirical covariance in cases of low number of observations to the number of assets in the portfolio.

4.2.2 Empirical data

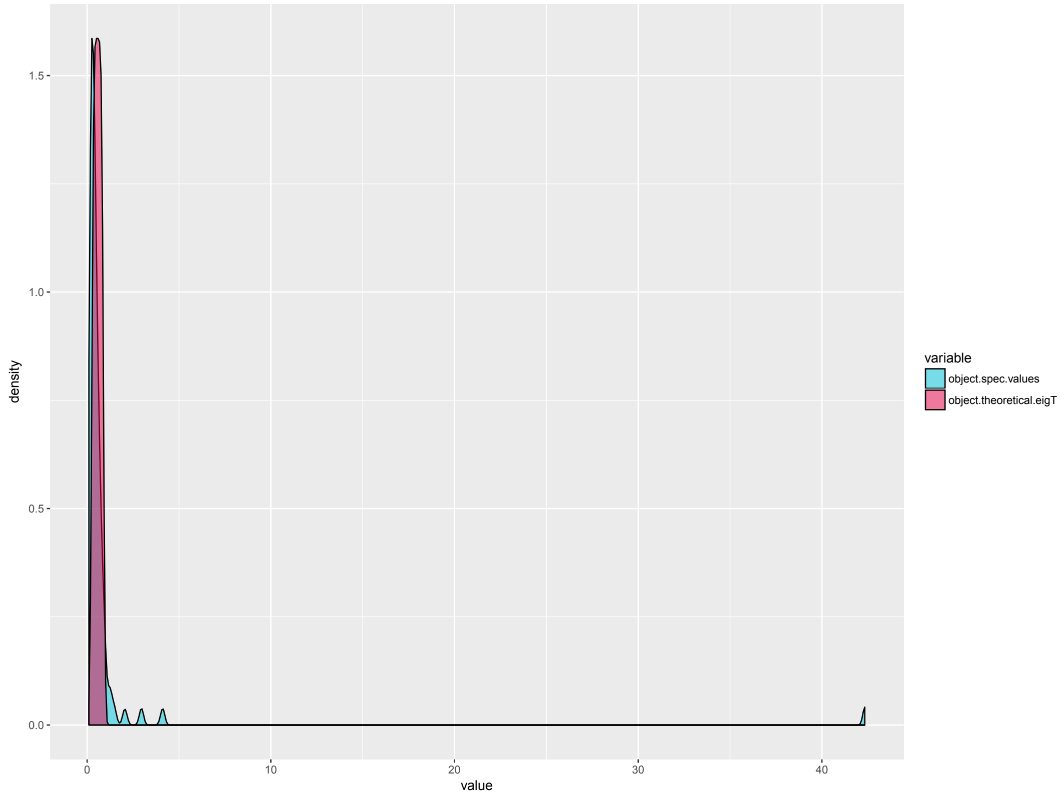
Applying the same procedure to the empirical data, same conclusion can be achieved. Empirical matrix diverges for low Qs and converges with the RMT estimate at $Q = 4$. For larger ratios, the estimates are practically same distance from the target matrix. Since in this case, the true covariance structure is unknown, the covariance matrix of the whole time series is taken as target. Both estimates converge to the target matrix. It is to be expected no later than for maximum Q, since it is how the target matrix is calculated and the calculation is deterministic. This also shows, that the Kullback-Leibler distance is a sensible measure in this instance.

Figure 4.2: K-L distances for empirical data



This shows that RMT is relevant particularly for noisy cases, which are approximated by the Q coefficient as the ratio of number of portfolio observations to the number of components. Therefore, under those conditions we expect the portfolios to be less risky in terms of Expected Shortfall, as the covariance matrix is cleansed and therefore closer to the 'true' one. We hypothesize that more precise information would allow the algorithm to make a better decision. Higher returns can be expected hand in hand with the portfolios. In other words, we might conjecture the Markowitz's Efficient frontier is to shift to the left and upwards. Although this is not within the scope of this thesis and is left for further research.

Figure 4.3: Eigenvalues spectrum of empirical covariance matrix



In Chapter 3 we presented the Marcenko-Pastur density of eigenvalues of the covariance matrix. Our data exhibit the behaviour described theoretically. The underlying data come from the original dataset of S&P 500 constituents and all observations are taken in to account. The spectrum coincides with the theoretical bounds set by the M-P band, while also having a number of eigenvalues greater than the upper bound of the M-P interval. There is total 11 eigenvalues outside the noise bound. Also, the market mode, or market trend, is represented by a disproportionally large eigenvalue, whose elements have a close to uniform distribution. It is almost 8 times larger than the second largest one. This result is in accord with others in the current literature.

4.3 Main results

The primary objective of this thesis is to determine, whether portfolios with RMT covariance matrices are less risky in terms of their Expected Shortfalls, under the assumption that returns follow Stable distributions. Or whether is it even better to directly minimise this risk measure instead. As in the previous part, we shall work with the Q coefficient, now in form or rebalancing. In all portfolios, there is a training period of size $N+1$, where N is the number of stocks – in our case 60. That means, that the portfolios are rebalanced by $60 * Q$ days with training period of 61 days, so that the covariance matrix can be estimated at all. Then for the next period, the weights from the previous one are used on unseen data and the performance of the portfolio is recorded for that period, until it is rebalanced again.

The risk exposure is measured parametrically by Expected Shortfall under the α-stable Lévy distributions. The formula to calculate is presented in previous chapter as Theorem 7 and is used to calculate both the ES and the respective VaR. We shall also report realised, or historical values for ES and VaR as a comparison between parametric and non-parametric approach.

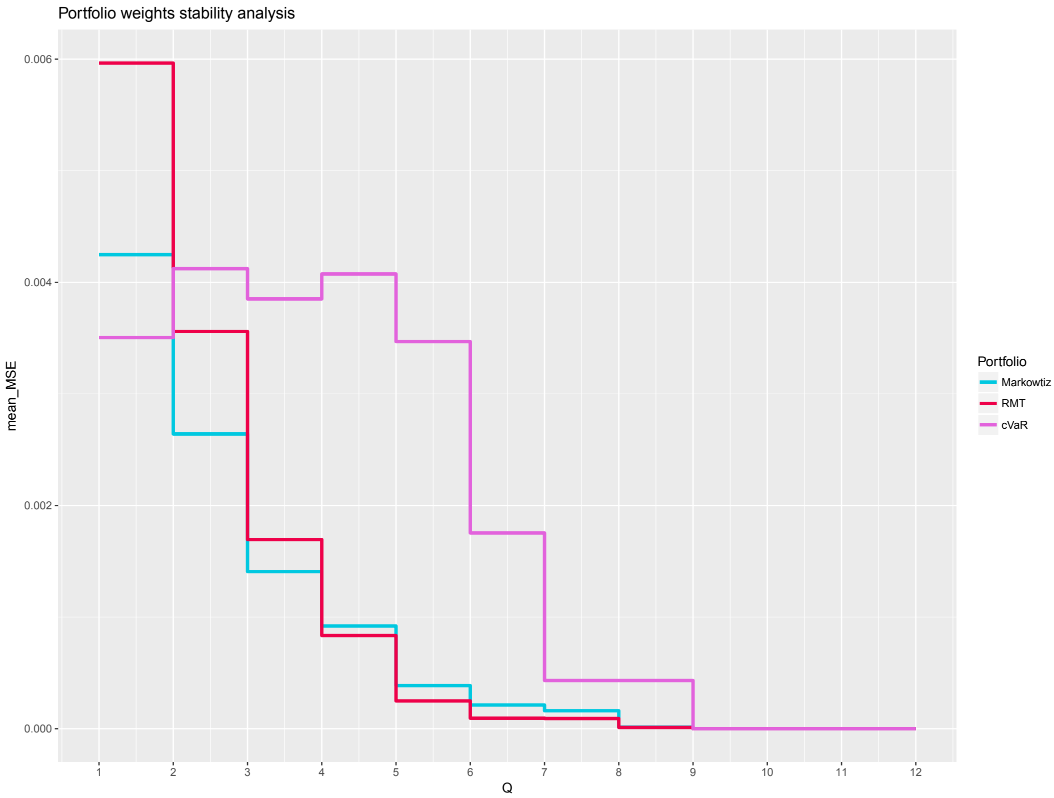
4.3.1 Rebalancing by Q

As stated in the original set of hypotheses, the influence of the Q is measured and inspected. In the discussion we shall consider three perspectives. The stability of individual portfolios, their reward and of course risk.

Weights analysis

We introduce a measure of weights stability as a divergence from those in the final portfolio, as it is the one with the most information available. The measure is the Mean Square Error with respect to the final weight of each portfolio constituent. In this case, the convergence is inspected with respect to Q. To obtain a single number for each period, the resulting index is a weighted average with respect to the weights in the portfolio.

Figure 4.4: Weights convergence with respect to Q

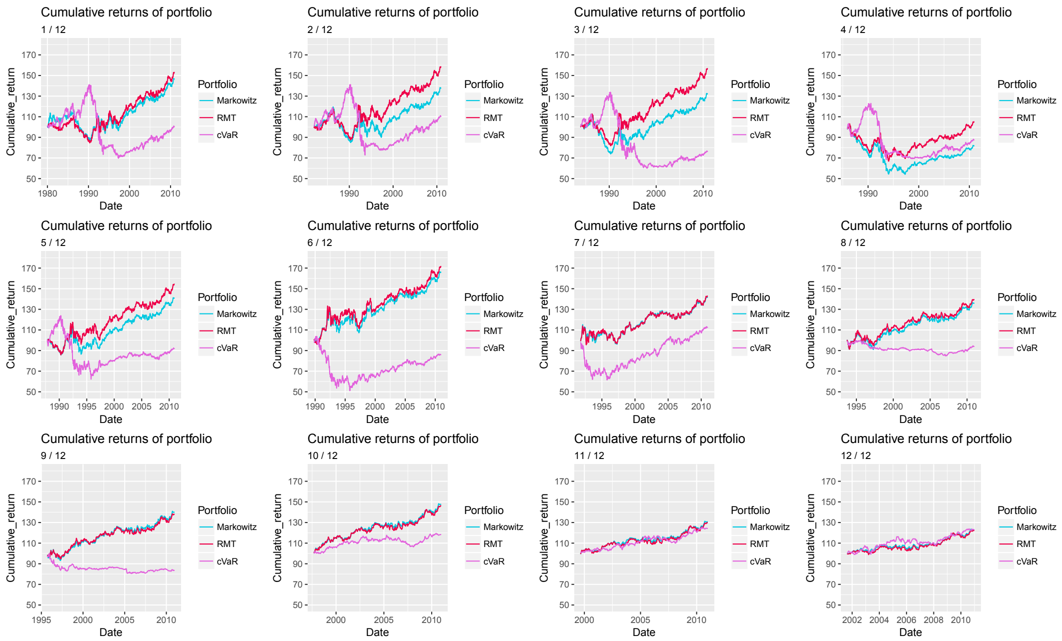


In case the portfolio was rebalanced by $Q = 1$, which in this case is 60 days, there is more error than for higher Q . Since it is mean of all the weights, for low rebalancing, there are more rebalancing period and as shown above, more noisy covariance matrices. Quite interestingly, the RMT and Markowitz portfolios show similar behaviour, whereas convergence of the cVaR portfolios is little slower. A closer inspection of the weights do not seem to offer a straight conclusion to this, but this is only to illustrate the main results. More detailed summary of weights such as mean or standard deviation is omitted, as they are not significantly different across the methods.

Returns analysis

Naturally, an important aspect of a portfolio is its return to the investor. In the settings of this work, this aspect yields very interesting conclusions. Since the period intentionally includes a market crash, we can compare how individual methods perform around it.

Figure 4.5: Returns of portfolios rebalanced by Q



We see, that cVaR portfolios react very badly to the credit crunch of 2008 – the data set by chance includes *AIG* stock – which plunges from price level of almost \$1000 to a penny stocks. Its price path is included in Appendix as Figure 1. Respectively, cVaR for $Q \in < 1, .., 4 >$ outperforms the other portfolios by quite a margin, but as the crisis come, the covariance–based methods are able to restructure very quickly and take advantage of shorting the worst stocks. cVaR does not seem to be able to do that at all. cVaR tends to go long little more often than take short positions and hence is more susceptible to market crashes.

Another reason can be the absence of the covariance matrix and therefore inability to incorporate the negative correlations in the optimisation. It is interesting to point out, that the whole S&P 500 index dips as low as onto level of 40 in the cumulative return. The index performance is depicted in the Appendix as Figure 2.

Notably, the Random Matrix Theory never performs worse than the Markowitz in terms of cumulative return. Despite their return paths being highly correlated. To inspect differences in returns and their significance, we employ simulation methods since in this case, we only see one realisation of the random process and the cumulative returns are heavily influenced by the order. Hence tests such as *t-test* for non-zero significant means little less relevant. Nevertheless, let us present them for rebalancing by $Q = 1$ and omit others, since some of the p-values only become significant for $Q = 11$ and $Q = 12$, and that on 10% confidence level.

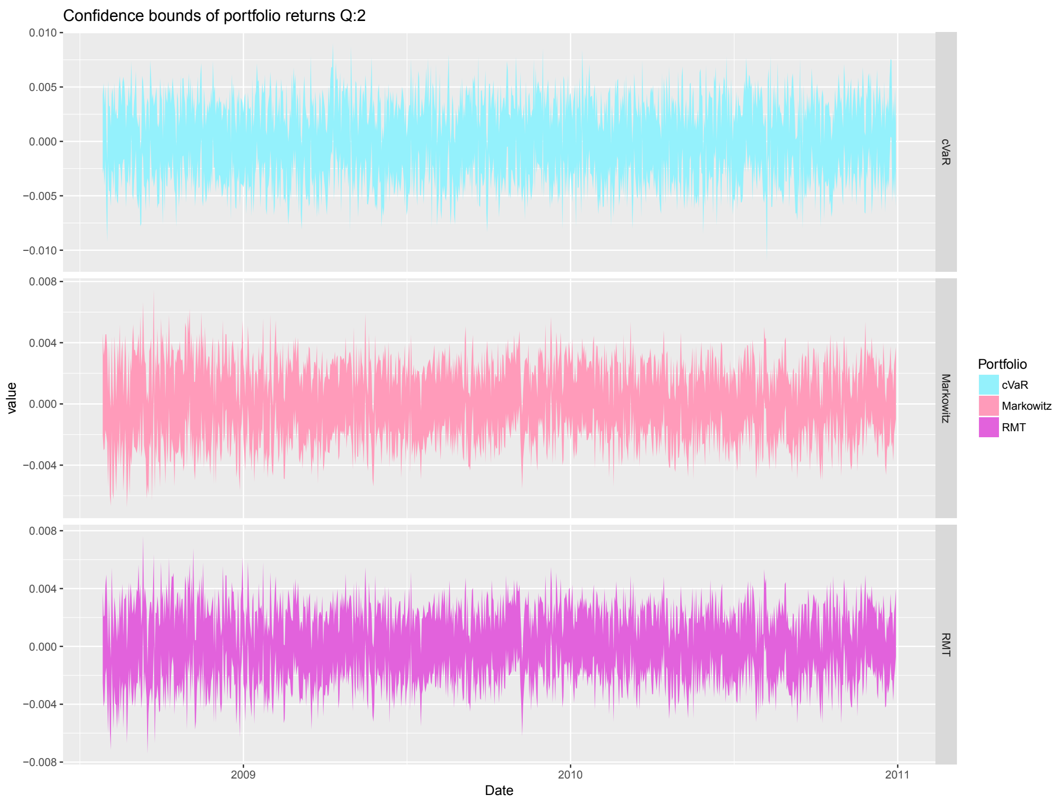
	Markowitz	RMT	cVaR
Markowitz	0.749		
RMT	0.658	0.758	
cVaR	0.889	0.599	0.669

Table 4.1: P-values of t-tests for Q-based rebalancing - Q 1 / 12

Simulated returns

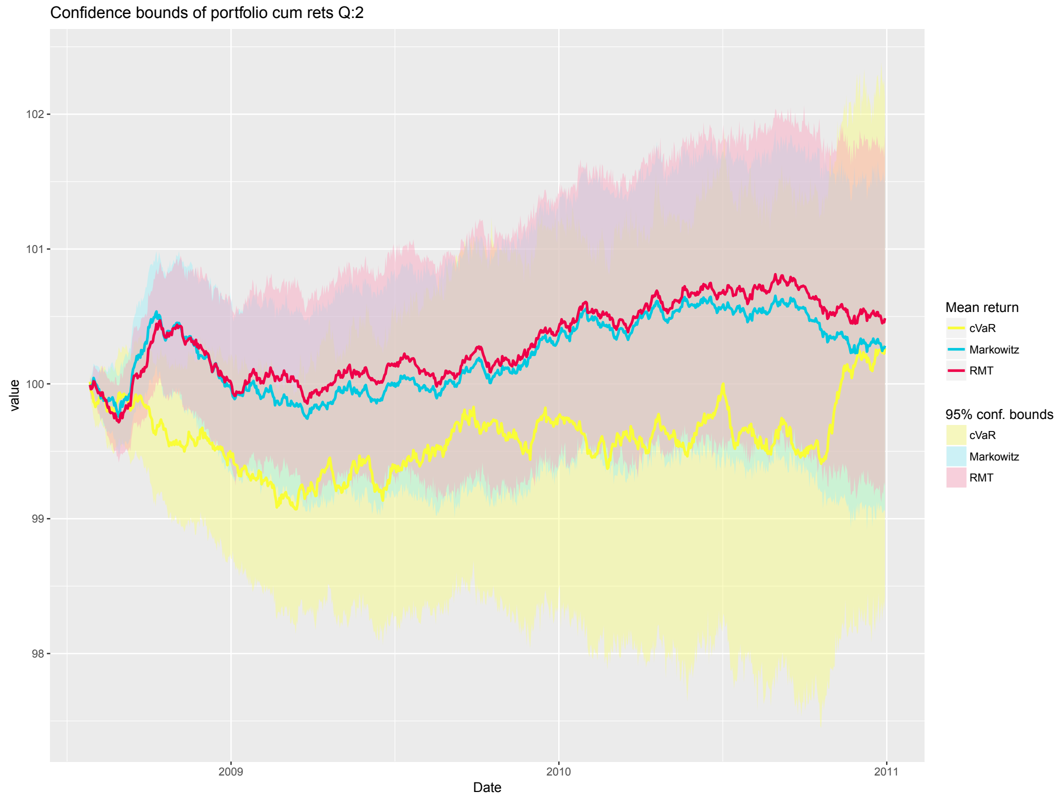
Therefore, we simulated 5000 realisations using the same covariance matrix to test whether or not the distributions are significantly different from zero. In terms of 95% confidence bands, for all portfolios they avoid the zero level in couple of subsequent blocks. As expected, the blocks are changing from being positive returns or negative returns and do not last very many days.

Figure 4.6: Simulated returns of portfolios rebalanced by Q



Let us present the case of $Q = 2$, as the other rebalancing periods do not yield much different results from this perspective. It is more interesting to look at the cumulative returns.

Figure 4.7: Simulated cumulative returns of portfolios rebalanced by Q



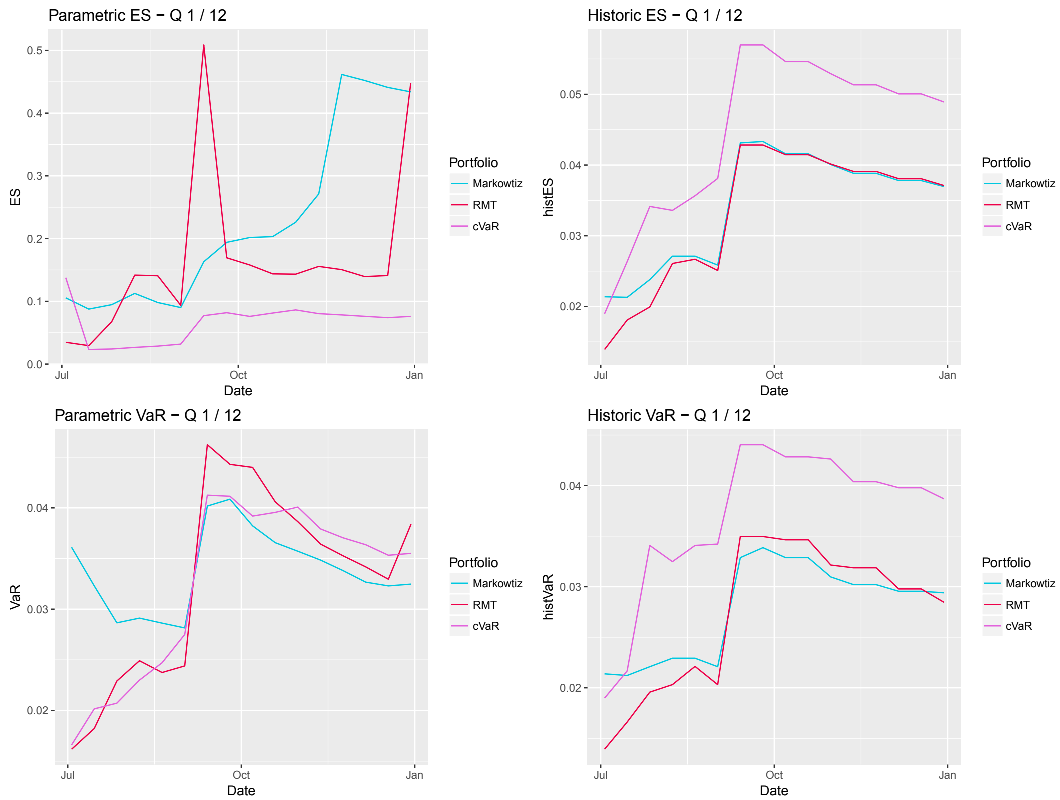
The confidence bounds widen over time as the variance of the underlying GMB also increases. The significance of cumulative return correlates with the actual returns in Figure 4.6 as the cVaR is for short periods of time significantly below its initial value. But not for long periods again.

The simulated results are heavily influenced by the nature of how they were created. GBM has time dependent volatility, but the volatility is not volatile itself, but rather constant in its magnitude. Its returns are mean centered and the portfolios therefore behave similarly. Hence, some form of stochastic volatility model, i.e. Heston model, might be better for this purpose or to use a real covariance matrix, rather than a simulated one. That is, however, left for further research.

Risk analysis

For each of the portfolio in the rebalancing period, we measure the parametric Expected shortfall attributed to the parameters from MLE fit. A side product of the ES calculation is the Value at Risk, which is necessary to measure losses beyond it. We also report historical figures for the ES and VaR as to compare the theoretical and realised risk over the holding period.

Figure 4.8: Risk of portfolios rebalanced by Q=1



When rebalancing by 60 days, with 61 days of training period, the cVaR portfolio, is from second period onwards the most stable and least risky in terms of the ES. However, the historic values are very correlated and cVaR portfolio is has higher ES than Markowitz's solution and the RMT portfolio. Since the realised values are very similar, the largest losses must be very similar and the distributions are differently shaped by losses below the VaR, because the parametric ES are quite different.

In terms of VaR[1], there are no large differences between the portfolios and in

[1]Note a different scale of the VaR, which is never greater than the ES, as it stems from the theoretical foundation.

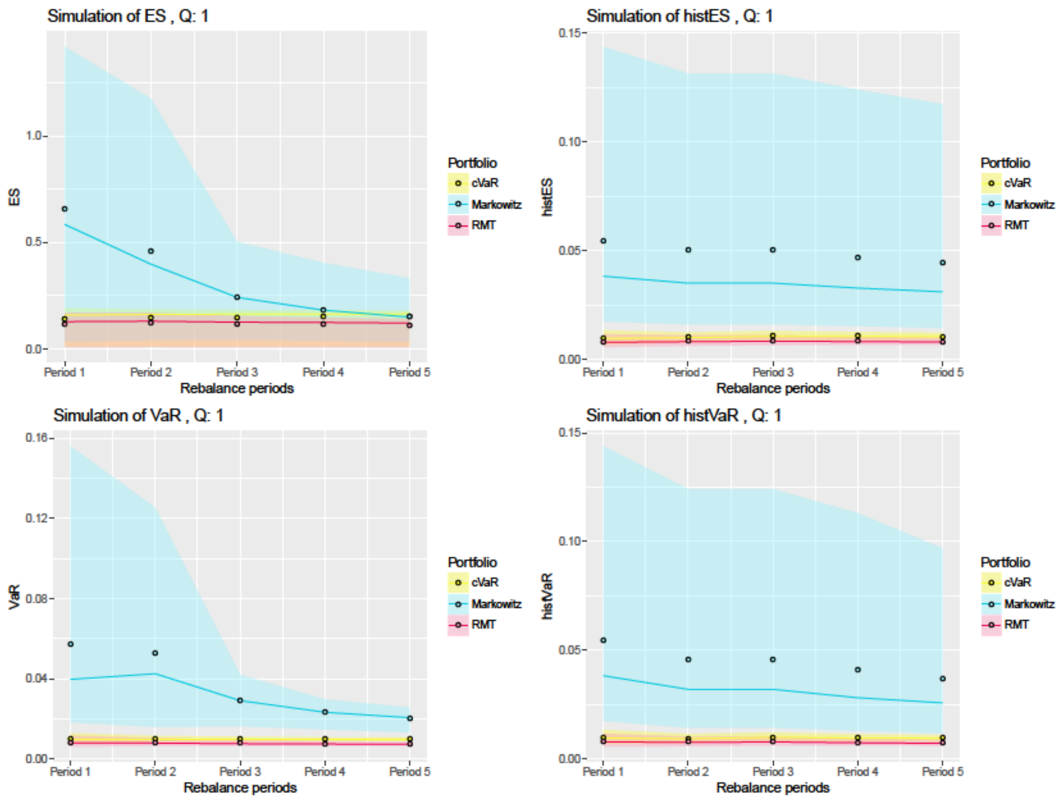
realised VaR there is a similar development to realised ES. Quite interestingly, the Markowitz and RMT values are very close, with RMT performing better in almost all of the cases. The high correlation can be expected given how are the portfolios computed. Moreover, the RMT yields lower risk portfolios at the beginning of the dataset, which was one of the hypothesis. Although, significance of the difference would need to be confirmed by simulations, since at this moment, there is no test how to determine it analytically.

Other rebalancing Qs are presented in the Appendix as Figures 3, 4 and 5 up to $Q = 4$ as , is the last to have at least 4 observations, since the rebalancing period gets too large. The results beyond are very similar to those presented here and in the Appendix.

Simulated risk

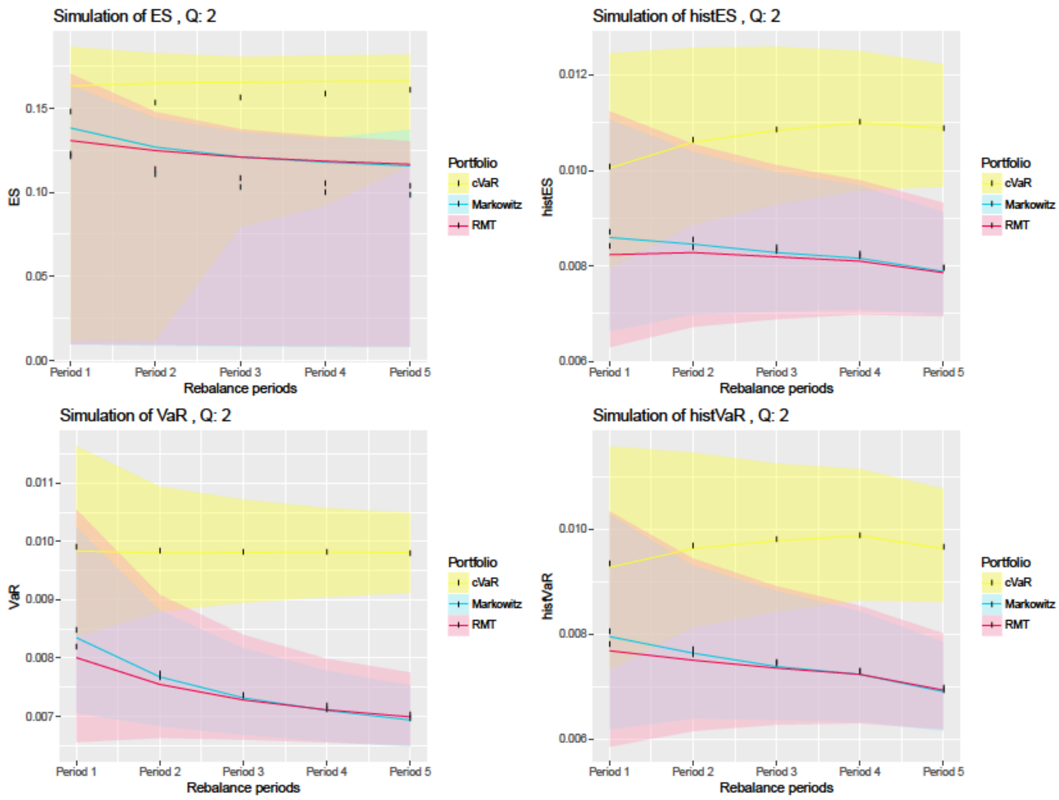
Using the same date of simulated returns, identical risk metrics were calculated so that the inference is done within constraints of a controlled environment. Since the calculation of the parametric ES and VaR, although much faster than the industry standard, is still very expensive, only two rebalancing parameters were considered and maximum number of rebalances limited to two. The MLE fit becomes quite computationally demanding for larger series, as the number of points in which density needs to be evaluated increases. In the following figures, the coloured points represent the mean of the simulations for the particular period and lines are medians of the same. The ribbons span the empirical 95% confidence interval obtained from the Monte Carlo simulations. Moreover, please note that the results are multiplied by a constant to enhance readability.

Figure 4.9: Risk of simulated portfolios rebalanced by Q=1



When rebalancing by $Q = 1$ in a controlled environment, the Markowitz's solution creates more risky portfolios in all of the measured quantities. In both cases, the mean and median are above those of cVaR and RMT portfolios. Markowitz's portfolios have also much larger variance. However, in parametric estimates, the Markowitz's portfolios converge quite quickly to RMT and cVaR. Considering that medians are lower than means, the distribution in individual periods are positively skewed towards zero.

Figure 4.10: Risk of simulated portfolios rebalanced by Q=2



Unlike the previous figure, when rebalancing in a controlled environment by $Q = 2$, the Markowitz's solution yields same portfolios with practically identical risk measures. The most volatile and risky portfolio is cVaR, which correlates with the cumulative returns depicted in Figure 4.7 where it starts with a plunge. That figure provides a sensible check to our computations, that if two portfolios have highly correlated performance, since the whole confidence bands of RMT and Markowitz are virtually identical the whole time, then their risk assessment is very similar.

Also, while cVaR portfolios' risks have a concave shape, the other two tend to decrease over the observed periods. This suggests that Markowitz's method and RMT better absorb the market information.

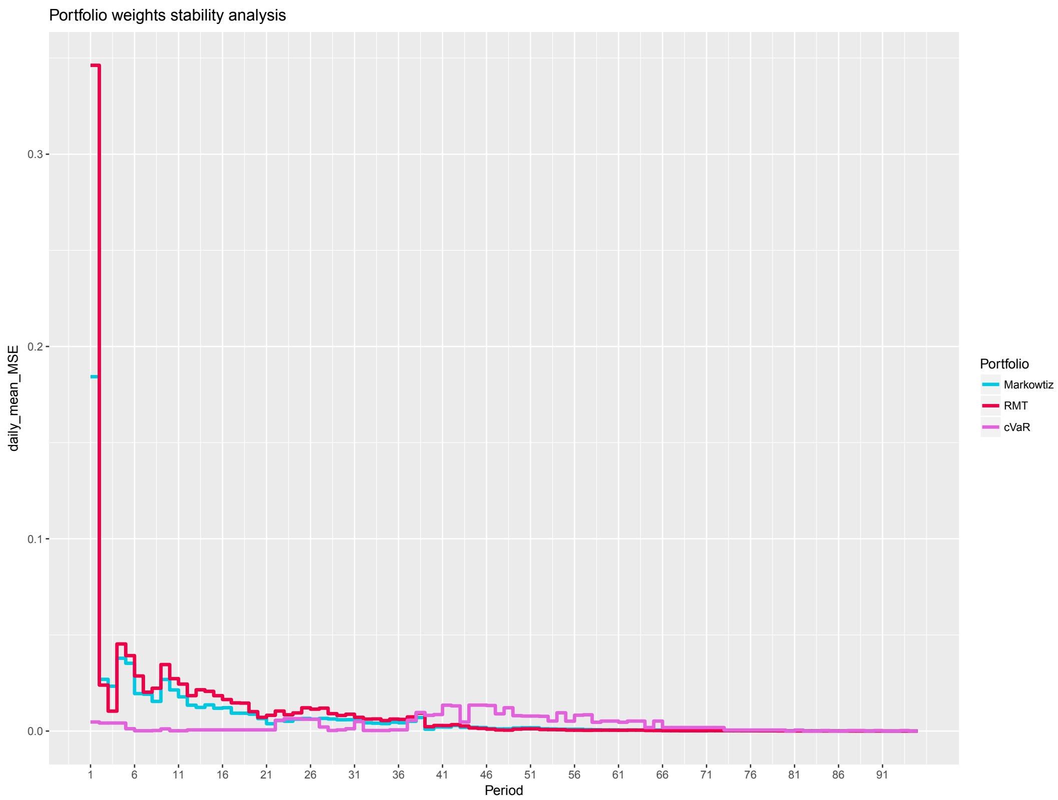
4.3.2 Rebalancing by days

Since the results of rebalancing by Q parameter, in fact a large time windows, show some interesting results with respect to the different behaviour of cVaR portfolios to RMT and Markowitz, we recalculated the results with respect to a shorter time interval of 10 days, or 2 trading weeks. The only difference to previous case is that now there is always a single set of results, as if the portfolios were rebalanced by $Q = 1/6$, but again with an initial training period of 61 days.

Weights analysis

The stability is calculated exactly same as in the previous case. This time the portfolios are only rebalanced one Q, as explained above.

Figure 4.11: Weights convergence with rebalancing by 10 days



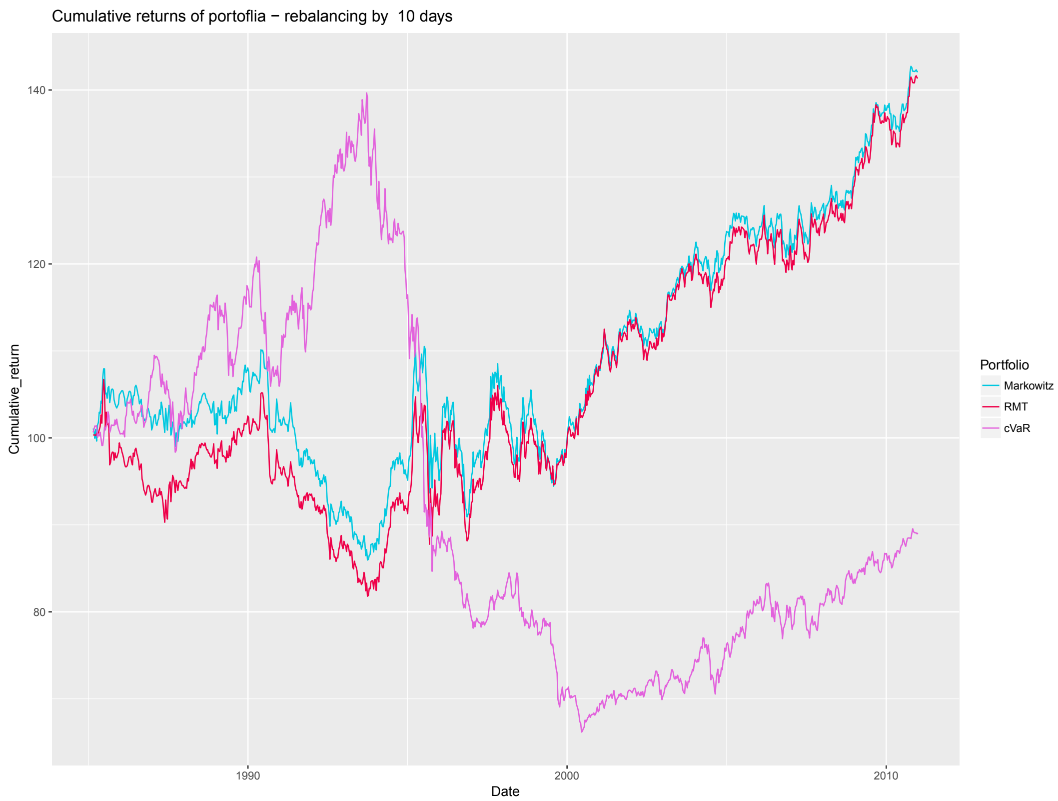
In case of shorter time interval for rebalancing, the cVaR portfolio quite notably stays within its scales from rebalancing by larger Q. Particularly in the beginning is quite close to its final weights, whereas other two methods only converge them after roughly 30 periods. But unlike them, in the middle of the period, where its returns drop significantly, the weights diverge a little in time of high volatility. The numerical results in this case show cVaR portfolio to be more stable than others.

Table 4.2: Comparison of weights stability

Statistic	N	Mean	St. Dev.	Min	Max
Markowtiz	135	0.006	0.010	0.000	0.066
RMT	135	0.007	0.012	0.000	0.094
cVaR	135	0.003	0.004	0.000	0.014

Returns analysis

Figure 4.12: Cumulative returns with rebalancing by 10 days

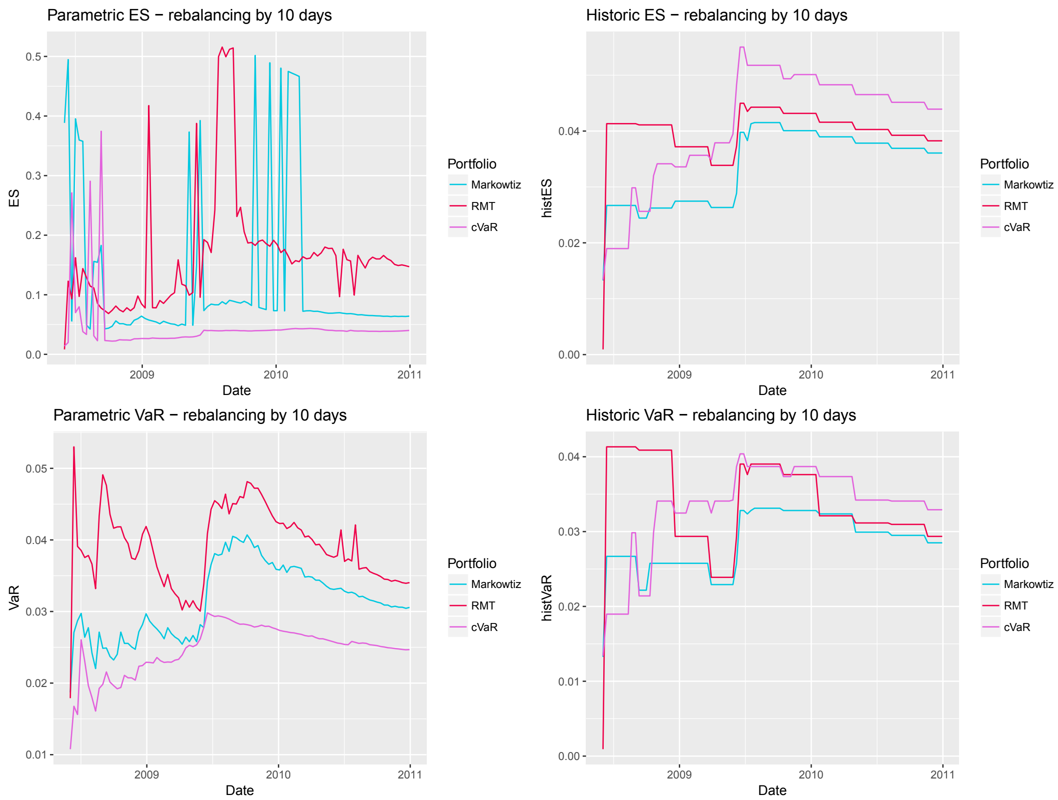


Similarly to the previous results, the cVaR portfolio is unable to avoid the dip around the time of the recession and does not manage to recover until the end of the observed period. Quite interestingly, the RMT underperforms against Markowitz at the beginning of the period, but from the recession onwards it caught up with Markowitz. The analysis of returns in terms of *t-test* and simulation results applies here.

Risk analysis

The risk is again measured as parametric value under the MLE fit of a Stable distribution in the same manner as in the previous case.

Figure 4.13: Cumulative returns with rebalancing by 10 days



The Markowitz and cVaR portfolios are quite volatile at the beginning. Regarding the Markowitz portfolio, it is something that was conjured as the covariance matrix estimate is noisy. Also, this portfolio attains its largest value in the second rebalancing period and couple of times again later. The cVaR solution is volatile and after a few periods remains stable as lower than the other two methods. RMT is quite lower and more stable at the beginning than Markowitz, which is also something that could be expected, however it does not avoid the occasional spike.

It is very interesting, that when comparing the parametric VaR with ES, the Markowitz's portfolios have always lower VaR, while it is not the case with the Expected Shortfall. In terms of the realised values, again, the performances are again highly correlated. Also just as in the case of rebalancing by Q, the cVaR portfolio is not significantly better than the other two.

4.4 Discussion

The most significant difference of cVaR method to RMT and Markowitz is the absence of any explicit covariance measure. The weights are calculated as a blackbox programming optimisation. Therefore they lack an intuitive component, unlike the other two. Following a closer inspection of the weights, the algorithm typically assigns only a relatively small set of values, so it is possible that the *Rglpk* engine only achieves a local optimum.Although, the problem of minimising the Expected shortfall is convex and smooth, it does not guarantee that the best solver is used and hence necessarily finds the global optimum. However, other solvers tried did not converge, or their interface did not allow for the necessary specification of the restrictions.

Also, the cVaR minimisation takes a minimum target return as a required input. Therefore it cannot find something such as Global Minimum Variance portfolio and is subjected to more uncertainty. For the purposes of this thesis, the target was always specified as 0. One of the possible extensions of this work would be to inspect the behaviour of the algorithms with target returns set and possibly inspect the effects on the Efficient Frontier.

cVaR tends to underperform in the initial periods. This suggests that it needs longer time series for the learning part to perform well. Although on average, in the simulated series, it in the end outperforms the other two methods, its span of returns is much wider than of the other two methods. It also adapts more slowly to a sudden change in market conditions.

Those remarks offer some ideas for further research to validate in controlled environment with Monte Carlo confidence intervals. Since as of now, there are no methods to compare the performance in terms of risk more thoroughly, except for simulation based methods.

The RMT in controlled environment did not outperform Markowitz, When rebalancing by $Q = 2$ they yielded highly similar results, although the empirical results of K-L divergence to the true covariance matrix suggested, it should have. The covariance matrix distances only start to convergence when $Q = 3$. It was shown that although the Pearsonian matrix is much noisier than RMT's covariance, both methods performed the same in terms of the riskiness of their portfolios. Possible extension of this would to simulate a number of covariance matrices and also take Monte Carlo confidence intervals to further inspect the differences in covariance matrix estimates.

Quite interestingly, when observing the differences in scales between parametric and realised risk metrics, Value at Risks are much more closer than the ES. This is interesting from the perspective of the VaR as single point value, compared to the ES which considers a whole tail of events. Supposing the distribution assumption is indeed correct, and the calculated VaR, and therefore ES, are also exactly calculated, as they would be realised in an infinite time. Then investor using only realised risk measures underestimates the real risk by almost an order of magnitude.

Also, the risk metrics of the simulations seem to contrast the empirical ones with respect to the volatility of the estimates. That is intuitively explainable, as the simulations report values averaged over hundreds of realisations.

Chapter 5

Conclusion

This thesis considers modern improvements of the standard portfolio selection problem, developed in the 1950's by the Nobel prize laureate, Harry Markowitz. Since the introduction of his solution to the optimal allocation of assets in the risk–return space, the problem has not witnessed much of an upheaval. Hence this work introduces two new methods of quite a different origin to test whether they could improve the original solution, particularly with respect to risk of the portfolios. The contribution of this work is theoretical as well as practical.

The topic departs from the very common assumption, that the returns of financial instruments are distributed normally, in the sense of Gaussian distributions. Instead, it replaces it with α-Lévy stable distributions, as was suggested by Benoit B. Mandelbrot in the 1960's, as an opposition to the Efficient Markets Hypothesis. This thesis offers a guide through sometimes confusing literature of this field. It also presents steps to simulate a financial market of arbitrary size with a known covariance matrix. This crucial part of the portfolio selection, is also the main subject of this work.

Special attention is paid to the Random Matrix Theory and the estimation of Pearsonian covariance matrix. The theory presents a method which denoises the Pearsonian covariance matrix and should therefore offer a better approximation of the true matrix. This is validated by re–estimating a known covariance matrix. For highly noised cases, the RMT performed better in simulated as well as empirical environment. The amount of noise is proxied by the ratio of number of observation to the number of observed assets.

The traditional risk measure in Markowitz framework, the Standard deviation, is replaced with Expected shortfall. This is motivated by its more suitable mathematical properties as a coherent risk measure and its regulatory necessity as of Basel III accord. Also, direct minimisation of the Expected shortfall is tested, as a portfolio management tool, against the other two methods.

Results do not suggest that Markowitz's solution is obsolete compared to the other two methods. Or rather, it did not perform significantly worse in terms of returns, or risk. Regarding returns, both in empirical and controlled environment, Markowitz–based portfolios highly correlate with RMT. Portfolios constructed by minimisation of the ES tend to underperform due to longer learning periods and lower flexibility in switching weights.

In terms of risk, the results are mixed in empirical cases. For the computational exhaustiveness, only a couple of rebalances could be calculated in sufficient number of simulations. Following the discussion of the results, possibly a different price generating engine than Geometric Brownian Motion could provide more conclusive results. RMT and Markowitz yield similar risk characteristics in the Monte Carlo simulations. With empirical stocks, the RMT has lower values of risk metrics on average, but results do not suggest an unambiguous conclusion.

The results also discovered that realised and parametric Value at Risk correlate very much even for periods with high order of noise. This thesis finds that although RMT should theoretically yield better estimates of the covariance matrix for low Qs, and subsequently 'better portfolios', this conjecture was not confirmed in full with Monte Carlo simulations.

Bibliography

ACERBI, C. (2002): "Spectral measures of risk: a coherent representation of subjective risk aversion." *Journal of Banking & Finance* **26(7)**: pp. 1505–1518.

ALEXANDER, S. S. (1961): "Price movements in speculative markets: Trends or random walks." *Industrial Management Review* **2**: pp. 7–26.

AROUS, G. B. & A. GUIONNET (2008): "The spectrum of heavy tailed random matrices." *Communications in Mathematical Physics* **278(3)**: pp. 715–751.

ARTZNER, P., F. DELBAEN, J.-M. EBER, & D. HEATH (1999): "Coherent measures of risk." *Mathematical finance* **9(3)**: pp. 203–228.

BACHELIER, L. (1900): *Théorie de la spéculation.* Gauthier-Villars.

BATES, D. & M. MAECHLER (2016): *Matrix: Sparse and Dense Matrix Classes and Methods.* R package version 1.2-6.

BERAN, J. (1994): *Statistics for Long-Memory Processes.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

BORAK, S., W. HÄRDLE, & R. WERON (2005): "Stable distributions." In "Statistical tools for finance and insurance," pp. 21–44. Springer.

BOUCHAUD, J.-P. & M. POTTERS (2003): *Theory of financial risk and derivative pricing: from statistical physics to risk management.* Cambridge university press.

BOUCHAUD, J.-P. & M. POTTERS (2009): "Financial applications of random matrix theory: a short review." *arXiv preprint arXiv:0910.1205* .

BUN, J., J.-P. BOUCHAUD, & M. POTTERS (2017): "Cleaning large correlation matrices: tools from random matrix theory." *Physics Reports* **666**: pp. 1–109.

Davis, M., L. Bachelier, A. Etheridge, & P. Samuelson (2011): *Louis Bachelier's Theory of Speculation: The Origins of Modern Finance*. Princeton University Press.

Descartes, R. (2017): *The principles of philosophy*. BookRix.

Dickinson, J. P. (1974): "The reliability of estimation procedures in portfolio analysis." *Journal of Financial and Quantitative Analysis* **9(03)**: pp. 447–462.

Dowd, K. (2007): *Measuring market risk*. John Wiley & Sons.

Duchi, J. (2007): "Derivations for linear algebra and optimization." *Berkeley, California* .

Eddelbuettel, D. & R. Francois (2011): *Rcpp: Seamless R and C++ Integration*. R package version 3011.87.

Edelman, A. & N. R. Rao (2005): "Random matrix theory." *Acta Numerica* **14**: pp. 233–297.

Elton, E. J. & M. J. Gruber (1997): "Modern portfolio theory, 1950 to date." *Journal of Banking & Finance* **21(11)**: pp. 1743–1759.

Elton, E. J., M. J. Gruber, S. J. Brown, & W. N. Goetzmann (2009): *Modern portfolio theory and investment analysis*. John Wiley & Sons.

Frankfurter, G. M., H. E. Phillips, & J. P. Seagle (1971): "Portfolio selection: the effects of uncertain means, variances, and covariances." *Journal of Financial and Quantitative Analysis* **6(05)**: pp. 1251–1262.

Frey, R. & A. J. McNeil (2002): "Var and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights." *Journal of banking & finance* **26(7)**: pp. 1317–1334.

Gençay, R., F. Selçuk, & B. J. Whitcher (2001): *An introduction to wavelets and other filtering methods in finance and economics*. Academic press.

Gnedenko, B. V. & A. N. Kolmogorov (1954): *Limit distributions for sums of independent random variables*. Addison-Wesley Mathematics Series. Cambridge, MA: Addison-Wesley. Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob. MR:0062975. Zbl:0056.36001.

Higham, N. J. (2002): "Computing the nearest correlation matrix—a problem from finance." *IMA journal of Numerical Analysis* **22(3)**: pp. 329–343.

Hirschberger, M., Y. Qi, & R. E. Steuer (2007): "Randomly generating portfolio-selection covariance matrices with specified distributional characteristics." *European Journal of Operational Research* **177(3)**: pp. 1610–1625.

Hurst, H. E. (1951): "Long-term storage capacity of reservoirs." *Trans. Amer. Soc. Civil Eng.* **116**: pp. 770–808.

Johnstone, I. M. (2001): "On the distribution of the largest eigenvalue in principal components analysis." *Ann. Statist.* **29(2)**: pp. 295–327.

Jorion, P. (1986): "Bayes-stein estimation for portfolio analysis." *Journal of Financial and Quantitative Analysis* **21(03)**: pp. 279–292.

Kendall, M. G. & A. B. Hill (1953): "The analysis of economic time-series-part i: Prices." *Journal of the Royal Statistical Society. Series A (General)* **116(1)**: pp. 11–34.

Kondor, I., S. Pafka, & G. Nagy (2007): "Noise sensitivity of portfolio selection under various risk measures." *Journal of Banking & Finance* **31(5)**: pp. 1545–1573.

Kraus, A. & R. H. Litzenberger (1976): "Skewness preference and the valuation of risk assets." *The Journal of Finance* **31(4)**: pp. 1085–1100.

Kristoufek, L. & M. Vosvrda (2013): "Measuring capital market efficiency: Global and local correlations structure." *Physica A: Statistical Mechanics and its Applications* **392(1)**: pp. 184–193.

Laloux, L., P. Cizeau, J.-P. Bouchaud, & M. Potters (1999): "Noise dressing of financial correlation matrices." *Physical review letters* **83(7)**: p. 1467.

Laloux, L., P. Cizeau, M. Potters, & J.-P. Bouchaud (2000): "Random matrix theory and financial correlations." *International Journal of Theoretical and Applied Finance* **3(03)**: pp. 391–397.

Lambert, P. & J. K. Lindsey (1999): "Analysing financial returns by using regression models based on non-symmetric stable distributions." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **48(3)**: pp. 409–424.

LARSON, A. B. (1960): "Measurement of a Random Process in Futures Prices." *Food Research Institute Studies* .

LEE, C. F. (1977): "Functional form, skewness effect, and the risk-return relationship." *Journal of financial and quantitative analysis* **12(01)**: pp. 55–72.

LÉVY, P. (1925): *Calcul des probabilités*. Gauthier-Villars.

MALKIEL, B. G. & E. F. FAMA (1970): "Efficient capital markets: A review of theory and empirical work*." *The journal of Finance* **25(2)**: pp. 383–417.

MANDELBROT, B. (1963a): "New methods in statistical economics." *Journal of Political Economy* **71(5)**: pp. pp. 421–440.

MANDELBROT, B. (1963b): "The variation of certain speculative prices." *The Journal of Business* **36(4)**: pp. pp. 394–419.

MANDELBROT, B. (1967): "The variation of some other speculative prices." *Journal of Business* pp. 393–413.

MANDELBROT, B. & R. L. HUDSON (2007): *The Misbehavior of Markets: A fractal view of financial turbulence*. Basic books.

MARČENKO, V. A. & L. A. PASTUR (1967): "Distribution of eigenvalues for some sets of random matrices." *Mathematics of the USSR-Sbornik* **1(4)**: p. 457.

MARKOWITZ, H. (1952): "Portfolio selection." *The journal of finance* **7(1)**: pp. 77–91.

MARKOWITZ, H. (1990): "Foundations of portfolio theory. nobel lectures in economic sciences 1981-1990. ed. karl-goran maler."

MAUSSER, H. & D. ROSEN (2000): "Managing risk with expected shortfall." *Probabilistic Constrained Optimization: Methodology and Applications* **49**: p. 198.

MCCULLOCH, J. H. (1986): "Simple consistent estimators of stable distribution parameters." *Communications in Statistics-Simulation and Computation* **15(4)**: pp. 1109–1136.

MERTON, R. C. (1973): "An intertemporal capital asset pricing model." *Econometrica* **41(5)**: pp. 867–887.

Merton, R. C. (1980): "On estimating the expected return on the market: An exploratory investigation." *Journal of financial economics* **8(4)**: pp. 323–361.

Nolan, J. P. (1997): "Numerical calculation of stable densities and distribution functions." *Communications in statistics. Stochastic models* **13(4)**: pp. 759–774.

Nolan, J. P. (2001): "Maximum likelihood estimation and diagnostics for stable distributions." In "Lévy processes," pp. 379–400. Springer.

Nolan, J. P. (2003): "Modeling financial data with stable distributions." *Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book* **1**: pp. 105–130.

Nolan, J. P. (2012): *Stable distributions*, volume 1177108605. ISBN.

Osborne, M. F. (1959): "Brownian motion in the stock market." *Operations research* **7(2)**: pp. 145–173.

Pafka, S. & I. Kondor (2003): "Noisy covariance matrices and portfolio optimization ii." *Physica A: Statistical Mechanics and its Applications* **319**: pp. 487–494.

Pearson, K. (1905): "The problem of the random walk." *Nature* **72(1865)**: p. 294.

Plerou, V., P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, & H. E. Stanley (2002): "Random matrix approach to cross correlations in financial data." *Phys. Rev. E* **65**: p. 066126.

Plerou, V., P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, & H. E. Stanley (1999): "Universal and nonuniversal properties of cross correlations in financial time series." *Physical Review Letters* **83(7)**: p. 1471.

Potters, M., J.-P. Bouchaud, & L. Laloux (2005): "Financial applications of random matrix theory: Old laces and new pieces." *arXiv preprint physics/0507111* .

Rachev, S. & S. Mittnik (2000): "Stable paretian models in finance." *Series in Financial Economics and Quantitative Analysis* .

Rockafellar, R. T. (2015): *Convex analysis.* Princeton university press.

Rockafellar, R. T. & S. Uryasev (2000): "Optimization of conditional value-at-risk." *Journal of risk* **2**: pp. 21–42.

Rockafellar, R. T. & S. Uryasev (2002): "Conditional value-at-risk for general loss distributions." *Journal of banking & finance* **26(7)**: pp. 1443–1471.

Rowe, B. L. Y. (2016): *tawny: Clean Covariance Matrices Using Random Matrix Theory and Shrinkage Estimators for Portfolio Optimization.* R package version 2.1.6.

Samoradnitsky, G. & M. S. Taqqu (1994): *Stable non-Gaussian random processes: stochastic models with infinite variance*, volume 1. CRC press.

Samuelson, P. A. (1965): "Proof that properly anticipated prices fluctuate randomly." *Industrial management review* **6(2)**: pp. 41–49.

Sewell, M. (2011): "History of the efficient market hypothesis." *RN* **11(04)**: p. 04.

Sharpe, W. F. (1964): "Capital asset prices: A theory of market equilibrium under conditions of risk." *The journal of finance* **19(3)**: pp. 425–442.

Shor, N. Z. (2012): *Minimization methods for non-differentiable functions*, volume 3. Springer Science & Business Media.

Stoyanov, S. V., G. Samorodnitsky, S. Rachev, & S. Ortobelli Lozza (2006): "Computing the portfolio conditional value-at-risk in the alpha-stable case." *Probability and Mathematical Statistics* **26(1)**: pp. 1–22.

Tola, V., F. Lillo, M. Gallegati, & R. N. Mantegna (2008): "Cluster analysis for portfolio optimization." *Journal of Economic Dynamics and Control* **32(1)**: pp. 235–258.

Tracy, C. A. & H. Widom (1994): "Level-spacing distributions and the airy kernel." *Communications in Mathematical Physics* **159(1)**: pp. 151–174.

Venables, W. N. & B. D. Ripley (2002): *Modern Applied Statistics with S.* New York: Springer, fourth edition.

Ville, J. (1939): "Etude critique de la notion de collectif." *Bull. Amer. Math. Soc. 45 (1939), 824 DOI: http://dx. doi. org/10.1090/S0002-9904-1939-07089-4 PII* pp. 0002–9904.

Wigner, E. P. (1951): "On the statistical distribution of the widths and spacings of nuclear resonance levels." In "Mathematical Proceedings of the Cambridge Philosophical Society," volume 47, pp. 790–798. Cambridge University Press.

Wuertz, D., M. Maechler, & R. core team members. (2016): *stabledist: Stable Distribution Functions*. R package version 0.7-1.

Wuertz, D., T. Setz, Y. Chalabi, & R. C. Team (2014): *fBasics: Rmetrics - Markets and Basic Statistics*. R package version 3011.87.

Yamai, Y., T. Yoshiba *et al.* (2002): "Comparative analyses of expected shortfall and value-at-risk: their estimation error, decomposition, and optimization." *Monetary and economic studies* **20(1)**: pp. 87–121.

Zolotarev, V. M. (1986): *One-dimensional stable distributions*, volume 65. American Mathematical Soc.

Table 1: Stocks description - pt1

Statistic	N	Mean	St. Dev.	Min	Max
AA	1,006	−0.001	0.038	−0.175	0.209
AAP	1,006	0.001	0.026	−0.197	0.130
ABT	1,006	0.0001	0.015	−0.096	0.092
ADB	1,006	−0.0003	0.027	−0.211	0.134
ADP	1,006	0.0002	0.017	−0.085	0.112
AEP	1,006	0.00001	0.018	−0.091	0.124
AFL	1,006	0.0003	0.040	−0.460	0.265
AGN	1,006	0.001	0.018	−0.106	0.090
AIG	1,006	−0.003	0.076	−0.936	0.507
AMA	1,006	−0.0002	0.026	−0.135	0.133
AON	1,006	0.0003	0.018	−0.148	0.108
APA	1,006	0.001	0.029	−0.201	0.193
APD	1,006	0.0003	0.023	−0.131	0.137
AVP	1,006	−0.00004	0.024	−0.167	0.160
AVY	1,006	−0.0003	0.024	−0.159	0.115
AXP	1,006	−0.0003	0.036	−0.194	0.188
BA	1,006	−0.0002	0.023	−0.080	0.144
BAC	1,006	−0.001	0.052	−0.342	0.302
BAX	1,006	0.0002	0.016	−0.142	0.081
BBT	1,006	−0.0003	0.036	−0.266	0.212
BBY	1,006	−0.0003	0.026	−0.160	0.165
BCR	1,006	0.0001	0.014	−0.087	0.068
BDX	1,006	0.0003	0.015	−0.088	0.092
BIG	1,006	0.0003	0.034	−0.297	0.200
BK.	1,006	−0.0002	0.037	−0.317	0.222
BLL	1,006	0.0005	0.020	−0.103	0.108
BMS	1,006	0.0001	0.020	−0.100	0.118
CA	1,006	0.0001	0.022	−0.137	0.179
CAG	1,006	−0.00003	0.015	−0.089	0.088
CB	1,006	0.0002	0.026	−0.220	0.189
CCL	1,006	−0.00001	0.029	−0.145	0.123

Table 2: Stocks description - pt2

Statistic	N	Mean	St. Dev.	Min	Max
CEL	1,006	0.00004	0.026	−0.168	0.171
CI	1,006	−0.0002	0.033	−0.242	0.211
CIN	1,006	−0.0002	0.027	−0.224	0.168
CL	1,006	0.0003	0.014	−0.071	0.100
CLF	1,006	0.001	0.053	−0.310	0.264
CLX	1,006	0.0001	0.013	−0.076	0.094
CMI	1,006	0.001	0.038	−0.192	0.199
CMS	1,006	0.0002	0.018	−0.106	0.102
CNP	1,006	0.0001	0.019	−0.124	0.125
COG	1,006	0.0003	0.035	−0.202	0.232
CSC	1,006	−0.0003	0.023	−0.177	0.129
CSX	1,006	0.001	0.028	−0.117	0.141
CTA	1,006	−0.0003	0.020	−0.132	0.102
CTL	1,006	0.0003	0.021	−0.141	0.163
CVX	1,006	0.0004	0.022	−0.133	0.189
DIS	1,006	0.0002	0.022	−0.102	0.148
ED	1,006	0.0002	0.013	−0.069	0.088
HRB	1,006	−0.001	0.026	−0.134	0.171
KO	1,006	0.0004	0.015	−0.091	0.130
MMM	1,006	0.0002	0.017	−0.090	0.094
MO	1,006	0.0004	0.015	−0.133	0.152
SCH	1,006	−0.00004	0.032	−0.164	0.179
T	1,006	0.00003	0.018	−0.080	0.151
VMC	1,006	−0.001	0.030	−0.112	0.169
WDC	1,006	0.0005	0.033	−0.129	0.175
WFC	1,006	−0.00004	0.044	−0.272	0.283
WHR	1,006	0.0002	0.032	−0.137	0.176
WMB	1,006	0.00004	0.031	−0.181	0.234
XRX	1,006	−0.0003	0.028	−0.207	0.165

Table 3: Comparison of weights stability

Statistic	N	Mean	St. Dev.	Min	Max

Figure 1: AIG price development

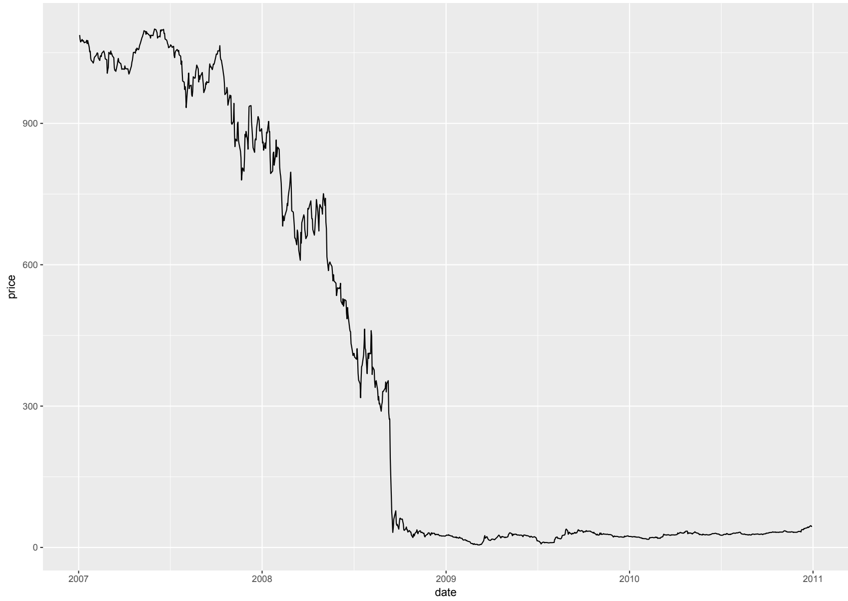


Figure 2: S&P 500 performance over the period

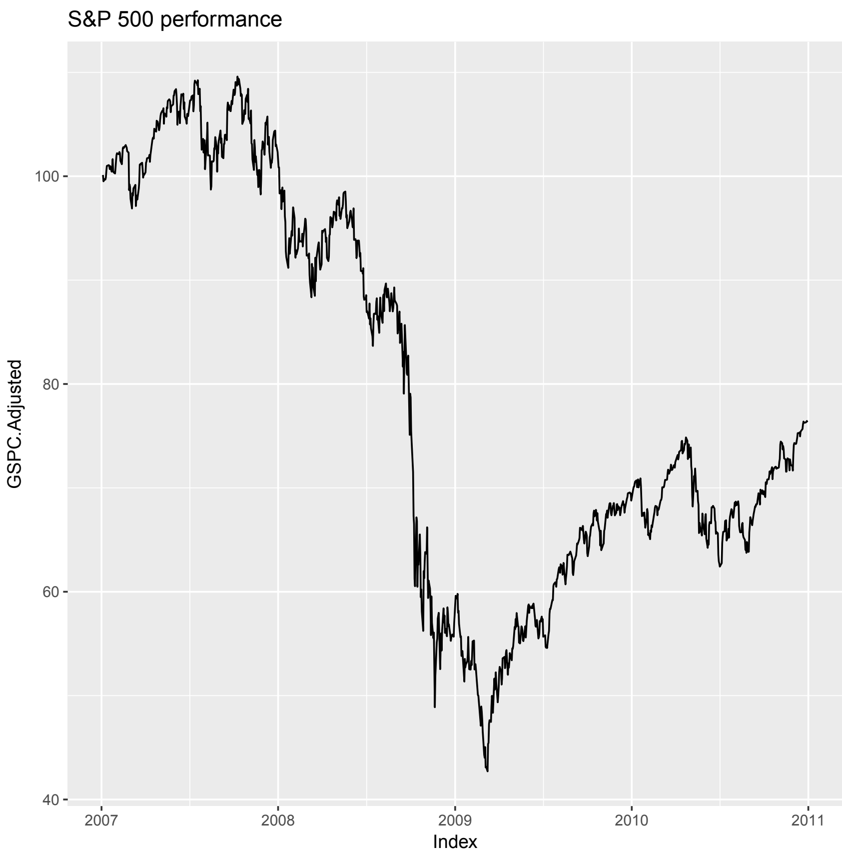


Figure 3: Risk of portfolios rebalanced by Q=2

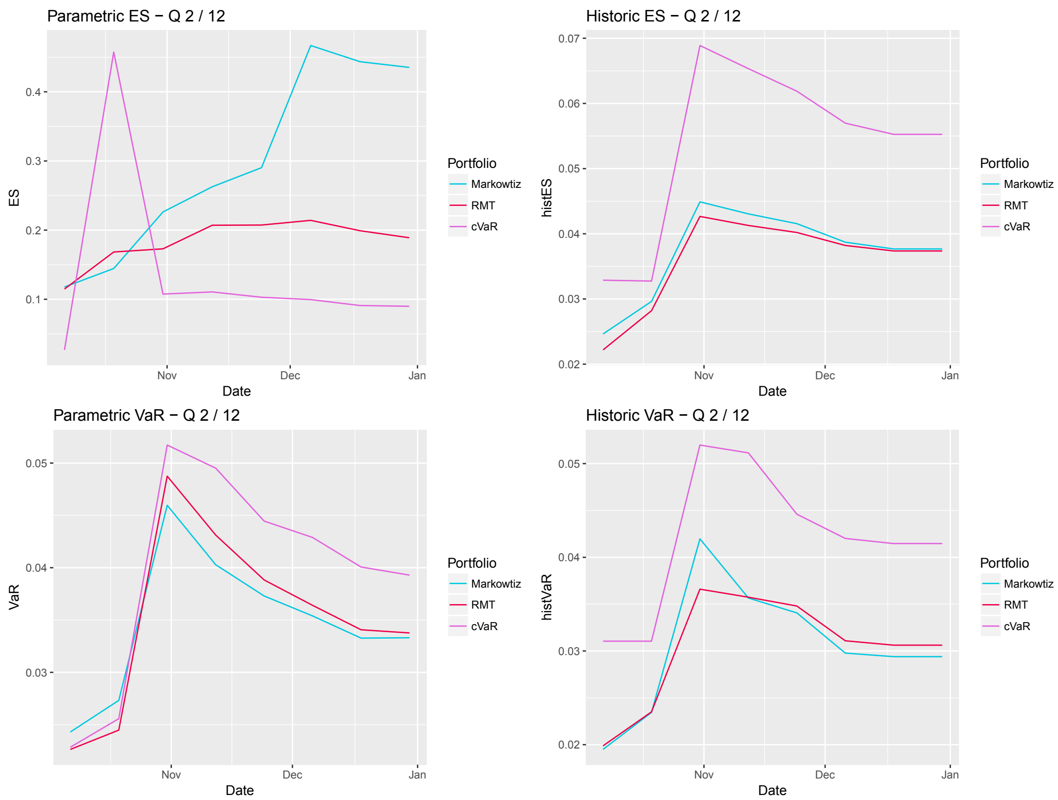


Figure 4: Risk of portfolios rebalanced by Q=3

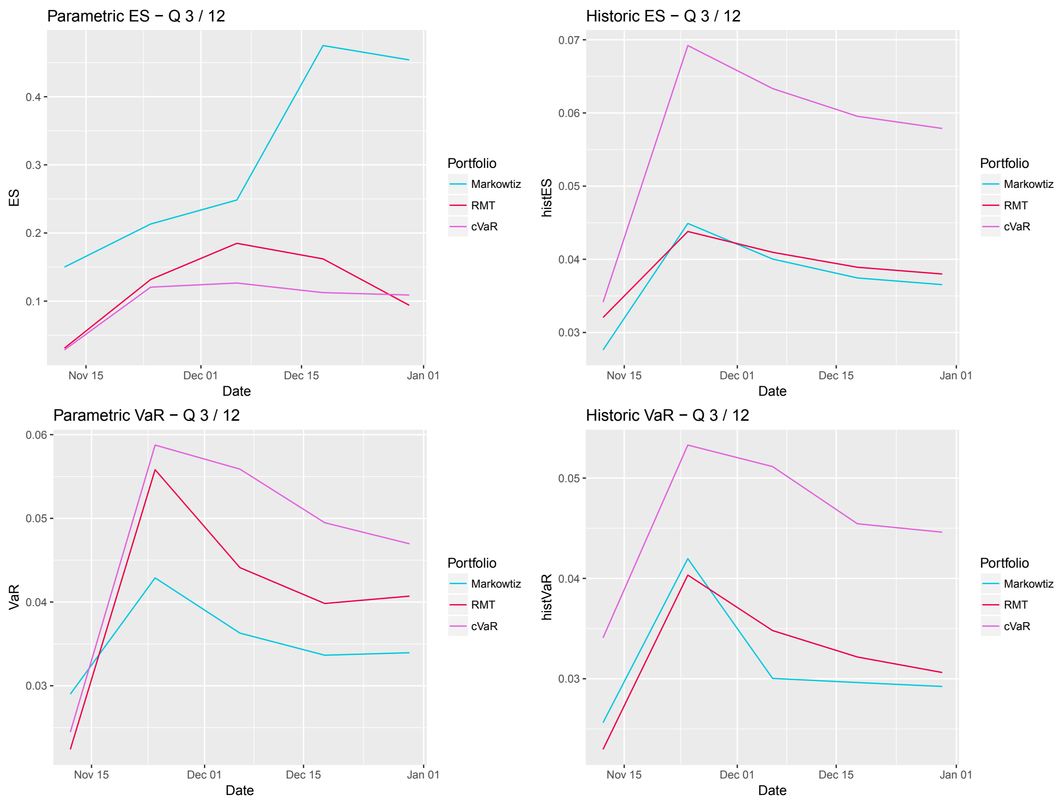


Figure 5: Risk of portfolios rebalanced by Q=4